(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 September 2002 (12.09.2002)

PCT

(10) International Publication Number
WO 02/070539 A2

(51) International Patent Classification[7]: C07K

(21) International Application Number: PCT/US02/05095

(22) International Filing Date: 5 March 2002 (05.03.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/799,451 5 March 2001 (05.03.2001) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 09/799,451 (CIP)
Filed on 5 March 2001 (05.03.2001)

(71) Applicant *(for all designated States except US)*: **HYSEQ, INC.** [US/US]; 670 Almanor Avenue, Sunnyvale, CA 94086 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **TANG, Y., Tom** [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **ZHOU, Ping** [US/US]; 7595 Newcastle Drive, Cupertino, CA 95014 (US). **GOODRICH, Ryle, W.** [US/US]; 4896 Sandy Lane, San Jose, CA 95124 (US). **ASUNDI, Vinod** [US/US]; 709 Foster City Boulevard, Foster City, CA 94404 (US). **ZHANG, Jie** [CN/US]; 4930 Poplar Terrace, Campbell, CA 95008 (US). **ZHAO, Qing, A.** [CN/US]; 1556 Kooser Road, San Jose, CA 95118 (US). **REN, Feiyan** [US/US]; 7703 Oak Meadow Court, Cupertino, CA 95014 (US). **XUE, Aidong, J.** [CN/US]; 1621 South Mary Avenue, Sunnyvale, CA 94087 (US). **YANG, Yonghong** [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **MA, Yunging** [CN/US]; 280 W. California Ave., #206, Sunnyvale, CA 94086 (US). **YAMAZAKI, Victoria** [JP/US]; 883 Portwalk Place, Redwood Shores, CA 94064 (US). **CHEN, Rui-hong** [US/US]; 1031 Flying Fish Street, Foster City, CA 94404 (US). **WANG, Zhiwei** [CN/US]; 836 Alturas Avenue, #B36, Sunnyvale, CA 94085 (US). **GHOSH, Malabika** [IN/US]; 620 Iris Avenue, Apt. 133, Sunnyvale, CA 94086 (US). **WEHRMAN, Tom** [US/US]; CCSR Mol Pharm 3210, 269 W. Campus Drive, Stanford, CA 94305 (US). **WANG, Jian-Rui** [CN/US]; 744 Stendahl Lane, Cupertino, CA 95014 (US). **WANG, Dunrui** [CN/US]; 12252 Pepper Tree Lane, Poway, CA 92064 (US). **DRMANAC, Radoje, T.** [US/US]; 850 East Greenwich Place, Palo Alto, CA 94303 (US).

(74) Agent: **ELRIFI, Ivor, R.**; Mintz, Levin, Cohn, Ferris, Glovsky and Popeo PC, One Fiancial Center, Boston, MA 02111 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
- *without international search report and to be republished upon receipt of that report*
- *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and uses thereof.

WO 02/070539 A2

NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

1. BACKGROUND OF THE INVENTION

1.1 TECHNICAL FIELD

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with uses for these polynucleotides and proteins, for example in therapeutic, diagnostic and research methods.

1.2 BACKGROUND

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs, chemokines, and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization-based cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity, for example, by virtue of their secreted nature in the case of leader sequence cloning, by virtue of their cell or tissue source in the case of PCR-based techniques, or by virtue of structural similarity to other genes of known biological activity.

Identified polynucleotide and polypeptide sequences have numerous applications in, for example, diagnostics, forensics, gene mapping; identification of mutations responsible for genetic disorders or other traits, to assess biodiversity, and to produce many other types of data and products dependent on DNA and amino acid sequences.

2. SUMMARY OF THE INVENTION

The compositions of the present invention include novel isolated polypeptides, novel isolated polynucleotides encoding such polypeptides, including recombinant DNA molecules, cloned genes or degenerate variants thereof, especially naturally occurring variants such as

allelic variants, antisense polynucleotide molecules, and antibodies that specifically recognize one or more epitopes present on such polypeptides, as well as hybridomas producing such antibodies.

The compositions of the present invention additionally include vectors, including expression vectors, containing the polynucleotides of the invention, cells genetically engineered to contain such polynucleotides and cells genetically engineered to express such polynucleotides.

The present invention relates to a collection or library of at least one novel nucleic acid sequence assembled from expressed sequence tags (ESTs) isolated mainly by sequencing by hybridization (SBH), and in some cases, sequences obtained from one or more public databases. The invention relates also to the proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins. These nucleic acid sequences are designated as SEQ ID NO: 1 – 948. The polypeptide sequences are designated SEQ ID NOS: 949-1896. The nucleic acids and polypeptides are provided in the Sequence Listing. In the nucleic acids provided in the Sequence Listing, A is adenine; C is cytosine; G is guanine; T is thymine; and N is any of the four bases. In the amino acids provided in the Sequence Listing, * corresponds to the stop codon.

The nucleic acid sequences of the present invention also include, nucleic acid sequences that hybridize to the complement of SEQ ID NO: 1 – 948 under stringent hybridization conditions; nucleic acid sequences which are allelic variants or species homologues of any of the nucleic acid sequences recited above, or nucleic acid sequences that encode a peptide comprising a specific domain or truncation of the peptides encoded by SEQ ID NO: 1 – 948. A polynucleotide comprising a nucleotide sequence having at least 90% identity to an identifying sequence of SEQ ID NO: 1 – 948 or a degenerate variant or fragment thereof. The identifying sequence can be 100 base pairs in length.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO: 1 – 948. The sequence information can be a segment of any one of SEQ ID NO: 1 – 948 that uniquely identifies or represents the sequence information of SEQ ID NO: 1 – 948.

A collection as used in this application can be a collection of only one polynucleotide. The collection of sequence information or identifying information of each sequence can be provided on a nucleic acid array. In one embodiment, segments of sequence information is provided on a nucleic acid array to detect the polynucleotide that contains the segment. The

array can be designed to detect full-match or mismatch to the polynucleotide that contains the segment. The collection can also be provided in a computer-readable format.

This invention also includes the reverse or direct complement of any of the nucleic acid sequences recited above; cloning or expression vectors containing the nucleic acid sequences; and host cells or organisms transformed with these expression vectors. Nucleic acid sequences (or their reverse or direct complements) according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology, such as use as hybridization probes, use as primers for PCR, use in an array, use in computer-readable media, use in sequencing full-length genes, use for chromosome and gene mapping, use in the recombinant production of protein, and use in the generation of anti-sense DNA or RNA, their chemical analogs and the like.

In a preferred embodiment, the nucleic acid sequences of SEQ ID NO: 1-948 or novel segments or parts of the nucleic acids of the invention are used as primers in expression assays that are well known in the art. In a particularly preferred embodiment, the nucleic acid sequences of SEQ ID NO: 1-948 or novel segments or parts of the nucleic acids provided herein are used in diagnostics for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The isolated polynucleotides of the invention include, but are not limited to, a polynucleotide comprising any one of the nucleotide sequences set forth in SEQ ID NO: 949-1896; a polynucleotide comprising any of the full length protein coding sequences of SEQ ID NO: 1–948; and a polynucleotide comprising any of the nucleotide sequences of the mature protein coding sequences of SEQ ID NO: 1–948. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent hybridization conditions to (a) the complement of any one of the nucleotide sequences set forth in SEQ ID NO: 1–948; (b) a nucleotide sequence encoding any one of the amino acid sequences set forth in the Sequence Listing; (c) a polynucleotide which is an allelic variant of any polynucleotides recited above; (d) a polynucleotide which encodes a species homolog (e.g. orthologs) of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of any of the polypeptides comprising an amino acid sequence set forth in the Sequence Listing.

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising any of the amino acid sequences set forth in the Sequence Listing; or the corresponding full length or mature protein. Polypeptides of the invention also include

polypeptides with biological activity that are encoded by (a) any of the polynucleotides having a nucleotide sequence set forth in SEQ ID NO: 1-948; or (b) polynucleotides that hybridize to the complement of the polynucleotides of (a) under stringent hybridization conditions. Biologically or immunologically active variants of any of the polypeptide sequences in the Sequence Listing, and "substantial equivalents" thereof (e.g., with at least about 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99% amino acid sequence identity) that preferably retain biological activity are also contemplated. The polypeptides of the invention may be wholly or partially chemically synthesized but are preferably produced by recombinant means using the genetically engineered cells (e.g. host cells) of the invention.

The invention also provides compositions comprising a polypeptide of the invention. Polypeptide compositions of the invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The invention also provides host cells transformed or transfected with a polynucleotide of the invention.

The invention also relates to methods for producing a polypeptide of the invention comprising growing a culture of the host cells of the invention in a suitable culture medium under conditions permitting expression of the desired polypeptide, and purifying the polypeptide from the culture or from the host cells. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Polynucleotides according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use as oligomers, or primers, for PCR, use for chromosome and gene mapping, use in the recombinant production of protein, and use in generation of anti-sense DNA or RNA, their chemical analogs and the like. For example, when the expression of an mRNA is largely restricted to a particular cell or tissue type, polynucleotides of the invention can be used as hybridization probes to detect the presence of the particular cell or tissue mRNA in a sample using, *e.g., in situ* hybridization.

In other exemplary embodiments, the polynucleotides are used in diagnostics as expressed sequence tags for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The polypeptides according to the invention can be used in a variety of conventional procedures and methods that are currently applied to other proteins. For example, a polypeptide of the invention can be used to generate an antibody that specifically binds the

polypeptide. Such antibodies, particularly monoclonal antibodies, are useful for detecting or quantitating the polypeptide in tissue. The polypeptides of the invention can also be used as molecular weight markers, and as a food supplement.

Methods are also provided for preventing, treating, or ameliorating a medical condition which comprises the step of administering to a mammalian subject a therapeutically effective amount of a composition comprising a polypeptide of the present invention and a pharmaceutically acceptable carrier.

In particular, the polypeptides and polynucleotides of the invention can be utilized, for example, in methods for the prevention and/or treatment of disorders involving aberrant protein expression or biological activity.

The present invention further relates to methods for detecting the presence of the polynucleotides or polypeptides of the invention in a sample. Such methods can, for example, be utilized as part of prognostic and diagnostic evaluation of disorders as recited herein and for the identification of subjects exhibiting a predisposition to such conditions. The invention provides a method for detecting the polynucleotides of the invention in a sample, comprising contacting the sample with a compound that binds to and forms a complex with the polynucleotide of interest for a period sufficient to form the complex and under conditions sufficient to form a complex and detecting the complex such that if a complex is detected, the polynucleotide of interest is detected. The invention also provides a method for detecting the polypeptides of the invention in a sample comprising contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex and detecting the formation of the complex such that if a complex is formed, the polypeptide is detected.

The invention also provides kits comprising polynucleotide probes and/or monoclonal antibodies, and optionally quantitative standards, for carrying out methods of the invention. Furthermore, the invention provides methods for evaluating the efficacy of drugs, and monitoring the progress of patients, involved in clinical trials for the treatment of disorders as recited above.

The invention also provides methods for the identification of compounds that modulate (i.e., increase or decrease) the expression or activity of the polynucleotides and/or polypeptides of the invention. Such methods can be utilized, for example, for the identification of compounds that can ameliorate symptoms of disorders as recited herein. Such methods can include, but are not limited to, assays for identifying compounds and other substances that interact with (*e.g.*, bind to) the polypeptides of the invention. The invention

provides a method for identifying a compound that binds to the polypeptides of the invention comprising contacting the compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and detecting the complex by detecting the reporter gene sequence expression such that if expression of the reporter gene is detected the compound the binds to a polypeptide of the invention is identified.

The methods of the invention also provides methods for treatment which involve the administration of the polynucleotides or polypeptides of the invention to individuals exhibiting symptoms or tendencies. In addition, the invention encompasses methods for treating diseases or disorders as recited herein comprising administering compounds and other substances that modulate the overall activity of the target gene products. Compounds and other substances can effect such modulation either on the level of target gene/protein expression or target protein activity.

The polypeptides of the present invention and the polynucleotides encoding them are also useful for the same functions known to one of skill in the art as the polypeptides and polynucleotides to which they have homology (set forth in Table 2); for which they have a signature region (as set forth in Table 3); or for which they have homology to a gene family (as set forth in Table 4). If no homology is set forth for a sequence, then the polypeptides and polynucleotides of the present invention are useful for a variety of applications, as described herein, including use in arrays for detection.

3. DETAILED DESCRIPTION OF THE INVENTION

3.1 DEFINITIONS

It must be noted that as used herein and in the appended claims, the singular forms "a", "an" and "the" include plural references unless the context clearly dictates otherwise.

The term "active" refers to those forms of the polypeptide which retain the biologic and/or immunologic activities of any naturally occurring polypeptide. According to the invention, the terms "biologically active" or "biological activity" refer to a protein or peptide having structural, regulatory or biochemical functions of a naturally occurring molecule. Likewise "immunologically active" or "immunological activity" refers to the capability of the natural, recombinant or synthetic polypeptide to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "activated cells" as used in this application are those cells which are engaged in extracellular or intracellular membrane trafficking, including the export of secretory or enzymatic molecules as part of a normal or disease process.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence 5'-AGT-3' binds to the complementary sequence 3'-TCA-5'. Complementarity between two single-stranded molecules may be "partial" such that only some of the nucleic acids bind or it may be "complete" such that total complementarity exists between the single stranded molecules. The degree of complementarity between the nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands.

The term "embryonic stem cells (ES)" refers to a cell that can give rise to many differentiated cell types in an embryo or an adult, including the germ cells. The term "germ line stem cells (GSCs)" refers to stem cells derived from primordial stem cells that provide a steady and continuous source of germ cells for the production of gametes. The term "primordial germ cells (PGCs)" refers to a small population of cells set aside from other cell lineages particularly from the yolk sac, mesenteries, or gonadal ridges during embryogenesis that have the potential to differentiate into germ cells and other cells. PGCs are the source from which GSCs and ES cells are derived The PGCs, the GSCs and the ES cells are capable of self-renewal. Thus these cells not only populate the germ line and give rise to a plurality of terminally differentiated cells that comprise the adult specialized organs, but are able to regenerate themselves.

The term "expression modulating fragment," EMF, means a series of nucleotides which modulates the expression of an operably linked ORF or another EMF.

As used herein, a sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of EMFs are nucleic acid fragments which induce the expression of an operably linked ORF in response to a specific regulatory factor or physiological event.

The terms "nucleotide sequence" or "nucleic acid" or "polynucleotide" or "oligonculeotide" are used interchangeably and refer to a heteropolymer of nucleotides or the sequence of these nucleotides. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA) or to any DNA-like or RNA-like material. In the sequences herein A is adenine, C is cytosine, T is thymine, G is guanine and

N is A, C, G or T (U). It is contemplated that where the polynucleotide is RNA, the T (thymine) in the sequences provided herein is substituted with U (uracil). Generally, nucleic acid segments provided by this invention may be assembled from fragments of the genome and short oligonucleotide linkers, or from a series of oligonucleotides, or from individual nucleotides, to provide a synthetic nucleic acid which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon, or a eukaryotic gene.

The terms "oligonucleotide fragment" or a "polynucleotide fragment", "portion," or "segment" or "probe" or "primer" are used interchangeably and refer to a sequence of nucleotide residues which are at least about 5 nucleotides, more preferably at least about 7 nucleotides, more preferably at least about 9 nucleotides, more preferably at least about 11 nucleotides and most preferably at least about 17 nucleotides. The fragment is preferably less than about 500 nucleotides, preferably less than about 200 nucleotides, more preferably less than about 100 nucleotides, more preferably less than about 50 nucleotides and most preferably less than 30 nucleotides. Preferably the probe is from about 6 nucleotides to about 200 nucleotides, preferably from about 15 to about 50 nucleotides, more preferably from about 17 to 30 nucleotides and most preferably from about 20 to 25 nucleotides. Preferably the fragments can be used in polymerase chain reaction (PCR), various hybridization procedures or microarray procedures to identify or amplify identical or related parts of mRNA or DNA molecules. A fragment or segment may uniquely identify each polynucleotide sequence of the present invention. Preferably the fragment comprises a sequence substantially similar to any one of SEQ ID NOs:1-948.

Probes may, for example, be used to determine whether specific mRNA molecules are present in a cell or tissue or to isolate similar nucleic acid sequences from chromosomal DNA as described by Walsh et al. (Walsh, P.S. et al., 1992, PCR Methods Appl 1:241-250). They may be labeled by nick translation, Klenow fill-in reaction, PCR, or other methods well known in the art. Probes of the present invention, their preparation and/or labeling are elaborated in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY; or Ausubel, F.M. et al., 1989, Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, both of which are incorporated herein by reference in their entirety.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NOs: 1-948. The sequence information can be a segment of any one of SEQ ID NOs: 1-948 that uniquely identifies or

represents the sequence information of that sequence of SEQ ID NO: 1-948. One such segment can be a twenty-mer nucleic acid sequence because the probability that a twenty-mer is fully matched in the human genome is 1 in 300. In the human genome, there are three billion base pairs in one set of chromosomes. Because 4^{20} possible twenty-mers exist, there are 300 times more twenty-mers than there are base pairs in a set of human chromosomes. Using the same analysis, the probability for a seventeen-mer to be fully matched in the human genome is approximately 1 in 5. When these segments are used in arrays for expression studies, fifteen-mer segments can be used. The probability that the fifteen-mer is fully matched in the expressed sequences is also approximately one in five because expressed sequences comprise less than approximately 5% of the entire genome sequence.

Similarly, when using sequence information for detecting a single mismatch, a segment can be a twenty-five mer. The probability that the twenty-five mer would appear in a human genome with a single mismatch is calculated by multiplying the probability for a full match ($1 \div 4^{25}$) times the increased probability for mismatch at each nucleotide position (3 x 25). The probability that an eighteen mer with a single mismatch can be detected in an array for expression studies is approximately one in five. The probability that a twenty-mer with a single mismatch can be detected in a human genome is approximately one in five.

The term "open reading frame," ORF, means a series of nucleotide triplets coding for amino acids without any termination codons and is a sequence translatable into protein.

The terms "operably linked" or "operably associated" refer to functionally related nucleic acid sequences. For example, a promoter is operably associated or operably linked with a coding sequence if the promoter controls the transcription of the coding sequence. While operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements e.g. repressor genes are not contiguously linked to the coding sequence but still control transcription/translation of the coding sequence.

The term "pluripotent" refers to the capability of a cell to differentiate into a number of differentiated cell types that are present in an adult organism. A pluripotent cell is restricted in its differentiation capability in comparison to a totipotent cell.

The terms "polypeptide" or "peptide" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide or protein sequence or fragment thereof and to naturally occurring or synthetic molecules. A polypeptide "fragment," "portion," or "segment" is a stretch of amino acid residues of at least about 5 amino acids, preferably at least about 7 amino acids, more preferably at least about 9 amino acids and most preferably at least about 17 or more amino acids. The peptide preferably is not greater than about 200 amino acids,

more preferably less than 150 amino acids and most preferably less than 100 amino acids. Preferably the peptide is from about 5 to about 200 amino acids. To be active, any polypeptide must have sufficient length to display biological and/or immunological activity.

The term "naturally occurring polypeptide" refers to polypeptides produced by cells that have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide including, but not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

The term "translated protein coding portion" means a sequence which encodes for the full length protein which may include any leader sequence or any processing sequence.

The term "mature protein coding sequence" means a sequence which encodes a peptide or protein without a signal or leader sequence. The "mature protein portion" means that portion of the protein which does not include a signal or leader sequence. The peptide may have been produced by processing in the cell which removes any leader/signal sequence. The mature protein portion may or may not include the initial methionine residue. The methionine residue may be removed from the protein during processing in the cell. The peptide may be produced synthetically or the protein may have been produced using a polynucleotide only encoding for the mature protein coding sequence.

The term "derivative" refers to polypeptides chemically modified by such techniques as ubiquitination, labeling (e.g., with radionuclides or various enzymes), covalent polymer attachment such as pegylation (derivatization with polyethylene glycol) and insertion or substitution by chemical synthesis of amino acids such as ornithine, which do not normally occur in human proteins.

The term "variant"(or "analog") refers to any polypeptide differing from naturally occurring polypeptides by amino acid insertions, deletions, and substitutions, created using, *e g.*, recombinant DNA techniques. Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest, may be found by comparing the sequence of the particular polypeptide with that of homologous peptides and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequence.

Alternatively, recombinant variants encoding these same or similar polypeptides may be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic system. Mutations in the polynucleotide sequence may be reflected

in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide, to change characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate.

Preferably, amino acid "substitutions" are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, *i.e.*, conservative amino acid replacements. "Conservative" amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Insertions" or "deletions" are preferably in the range of about 1 to 20 amino acids, more preferably 1 to 10 amino acids. The variation allowed may be experimentally determined by systematically making insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

Alternatively, where alteration of function is desired, insertions, deletions or non-conservative alterations can be engineered to produce altered polypeptides. Such alterations can, for example, alter one or more of the biological functions or biochemical characteristics of the polypeptides of the invention. For example, such alterations may change polypeptide characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate. Further, such alterations can be selected so as to generate polypeptides that are better suited for expression, scale up and the like in the host cells chosen for expression. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

The terms "purified" or "substantially purified" as used herein denotes that the indicated nucleic acid or polypeptide is present in the substantial absence of other biological macromolecules, *e.g.*, polynucleotides, proteins, and the like. In one embodiment, the polynucleotide or polypeptide is purified such that it constitutes at least 95% by weight, more preferably at least 99% by weight, of the indicated biological macromolecules present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 1000 daltons, can be present).

The term "isolated" as used herein refers to a nucleic acid or polypeptide separated from at least one other component (e.g., nucleic acid or polypeptide) present with the nucleic acid or polypeptide in its natural source. In one embodiment, the nucleic acid or polypeptide is found in the presence of (if anything) only a solvent, buffer, ion, or other component normally present in a solution of the same. The terms "isolated" and "purified" do not encompass nucleic acids or polypeptides present in their natural source.

The term "recombinant," when used herein to refer to a polypeptide or protein, means that a polypeptide or protein is derived from recombinant (*e.g.*, microbial, insect, or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (e.g., yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, *e.g., E. coli*, will be free of glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern in general different from those expressed in mammalian cells.

The term "recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. An expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an amino terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

The term "recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extrachromosomally. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed. This term also means host cells which have stably integrated a recombinant genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers. Recombinant expression systems as defined herein will express polypeptides or proteins

endogenous to the cell upon induction of the regulatory elements linked to the endogenous DNA segment or gene to be expressed. The cells can be prokaryotic or eukaryotic.

The term "secreted" includes a protein that is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence when it is expressed in a suitable host cell. "Secreted" proteins include without limitation proteins secreted wholly (*e.g.*, soluble proteins) or partially (*e.g.*, receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins that are transported across the membrane of the endoplasmic reticulum. "Secreted" proteins are also intended to include proteins containing non-typical signal sequences (e.g. Interleukin-1 Beta, see Krasney, P.A. and Young, P.R. (1992) Cytokine 4(2):134 -143) and factors released from damaged cells (e.g. Interleukin-1 Receptor Antagonist, see Arend, W.P. et. al. (1998) Annu. Rev. Immunol. 16:27-55)

Where desired, an expression vector may be designed to contain a "signal or leader sequence" which will direct the polypeptide through the membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous protein sources by recombinant DNA techniques.

The term "stringent" is used to refer to conditions that are commonly understood in the art as stringent. Stringent conditions can include highly stringent conditions (i.e., hybridization to filter-bound DNA in 0.5 M $NaHPO_4$, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1X SSC/0.1% SDS at 68°C), and moderately stringent conditions (i.e., washing in 0.2X SSC/0.1% SDS at 42°C). Other exemplary hybridization conditions are described herein in the examples.

In instances of hybridization of deoxyoligonucleotides, additional exemplary stringent hybridization conditions include washing in 6X SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligonucleotides), 48°C (for 17-base oligos), 55°C (for 20-base oligonucleotides), and 60°C (for 23-base oligonucleotides).

As used herein, "substantially equivalent" or "substantially similar" can refer both to nucleotide and amino acid sequences, for example a mutant sequence, that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between the reference and subject sequences. Typically, such a substantially equivalent sequence varies from one of those listed herein by no more than about 35% (*i.e.*, the number of individual residue substitutions, additions, and/or deletions in a substantially equivalent sequence, as compared to the

corresponding reference sequence, divided by the total number of residues in the substantially equivalent sequence is about 0.35 or less). Such a sequence is said to have 65% sequence identity to the listed sequence. In one embodiment, a substantially equivalent, *e.g.*, mutant, sequence of the invention varies from a listed sequence by no more than 30% (70% sequence identity); in a variation of this embodiment, by no more than 25% (75% sequence identity); and in a further variation of this embodiment, by no more than 20% (80% sequence identity) and in a further variation of this embodiment, by no more than 10% (90% sequence identity) and in a further variation of this embodiment, by no more that 5% (95% sequence identity). Substantially equivalent, *e.g.*, mutant, amino acid sequences according to the invention preferably have at least 80% sequence identity with a listed amino acid sequence, more preferably at least 85% sequence identity, more preferably at least 90% sequence identity, more preferably at least 95% sequence identity, more preferably at least 98% sequence identity, and most preferably at least 99% sequence identity. Substantially equivalent nucleotide sequence of the invention can have lower percent sequence identities, taking into account, for example, the redundancy or degeneracy of the genetic code. Preferably, the nucleotide sequence has at least about 65% identity, more preferably at least about 75% identity, more preferably at least about 80% sequence identity, more preferably at least 85% sequence identity, more preferably at least 90% sequence identity, more preferably at least about 95% sequence identity, more preferably at least 98% sequence identity, and most preferably at least 99% sequence identity. For the purposes of the present invention, sequences having substantially equivalent biological activity and substantially equivalent expression characteristics are considered substantially equivalent. For the purposes of determining equivalence, truncation of the mature sequence (*e.g.*, via a mutation which creates a spurious stop codon) should be disregarded. Sequence identity may be determined, e.g., using the Jotun Hein method (Hein, J. (1990) Methods Enzymol. 183:626-645). Identity between sequences can also be determined by other methods known in the art, e.g. by varying hybridization conditions.

The term "totipotent" refers to the capability of a cell to differentiate into all of the cell types of an adult organism.

The term "transformation" means introducing DNA into a suitable host cell so that the DNA is replicable, either as an extrachromosomal element, or by chromosomal integration. The term "transfection" refers to the taking up of an expression vector by a suitable host cell, whether or not any coding sequences are in fact expressed. The term "infection" refers to the introduction of nucleic acids into a suitable host cell by use of a virus or viral vector.

As used herein, an "uptake modulating fragment," UMF, means a series of nucleotides which mediate the uptake of a linked DNA fragment into a cell. UMFs can be readily identified using known UMFs as a target sequence or target motif with the computer-based systems described below. The presence and activity of a UMF can be confirmed by attaching the suspected UMF to a marker sequence. The resulting nucleic acid molecule is then incubated with an appropriate host under appropriate conditions and the uptake of the marker sequence is determined. As described above, a UMF will increase the frequency of uptake of a linked marker sequence.

Each of the above terms is meant to encompass all that is described for each, unless the context dictates otherwise.

3.2 NUCLEIC ACIDS OF THE INVENTION

Nucleotide sequences of the invention are set forth in the Sequence Listing.

The isolated polynucleotides of the invention include a polynucleotide comprising the nucleotide sequences of SEQ ID NO: 1 – 948; a polynucleotide encoding any one of the peptide sequences of SEQ ID NO: 949-1896; and a polynucleotide comprising the nucleotide sequence encoding the mature protein coding sequence of the polynucleotides of any one of SEQ ID NO: 1 - 948. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent conditions to (a) the complement of any of the nucleotides sequences of SEQ ID NO: 1 – 948; (b) nucleotide sequences encoding any one of the amino acid sequences set forth in the Sequence Listing as SEQ ID NO: 949-1896; (c) a polynucleotide which is an allelic variant of any polynucleotide recited above; (d) a polynucleotide which encodes a species homolog of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptides of SEQ ID NO: 949-1896. Domains of interest may depend on the nature of the encoded polypeptide; e.g., domains in receptor-like polypeptides include ligand-binding, extracellular, transmembrane, or cytoplasmic domains, or combinations thereof; domains in immunoglobulin-like proteins include the variable immunoglobulin-like domains; domains in enzyme-like polypeptides include catalytic and substrate binding domains; and domains in ligand polypeptides include receptor-binding domains.

The polynucleotides of the invention include naturally occurring or wholly or partially synthetic DNA, e.g., cDNA and genomic DNA, and RNA, e.g., mRNA. The polynucleotides may include all of the coding region of the cDNA or may represent a portion of the coding region of the cDNA.

The present invention also provides genes corresponding to the cDNA sequences disclosed herein. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. Further 5' and 3' sequence can be obtained using methods known in the art. For example, full length cDNA or genomic DNA that corresponds to any of the polynucleotides of SEQ ID NO: 1 – 948 can be obtained by screening appropriate cDNA or genomic DNA libraries under suitable hybridization conditions using any of the polynucleotides of SEQ ID NO: 1 - 948 or a portion thereof as a probe. Alternatively, the polynucleotides of SEQ ID NO: 1 - 948 may be used as the basis for suitable primer(s) that allow identification and/or amplification of genes in appropriate genomic DNA or cDNA libraries.

The nucleic acid sequences of the invention can be assembled from ESTs and sequences (including cDNA and genomic sequences) obtained from one or more public databases, such as dbEST, gbpri, and UniGene. The EST sequences can provide identifying sequence information, representative fragment or segment information, or novel segment information for the full-length gene.

The polynucleotides of the invention also provide polynucleotides including nucleotide sequences that are substantially equivalent to the polynucleotides recited above. Polynucleotides according to the invention can have, *e.g.*, at least about 65%, at least about 70%, at least about 75%, at least about 80%, 81%, 82%, 83%, 84%, more typically at least about 85%, 86%, 87%, 88%, 89%, more typically at least about 90%, 91%, 92%, 93%, 94%, and even more typically at least about 95%, 96%, 97%, 98%, 99% sequence identity to a polynucleotide recited above.

Included within the scope of the nucleic acid sequences of the invention are nucleic acid sequence fragments that hybridize under stringent conditions to any of the nucleotide sequences of SEQ ID NO: 1 - 948, or complements thereof, which fragment is greater than about 5 nucleotides, preferably 7 nucleotides, more preferably greater than 9 nucleotides and most preferably greater than 17 nucleotides. Fragments of, e.g. 15, 17, or 20 nucleotides or more that are selective for (i.e. specifically hybridize to any one of the polynucleotides of the invention) are contemplated. Probes capable of specifically hybridizing to a polynucleotide can differentiate polynucleotide sequences of the invention from other polynucleotide sequences in the same family of genes or can differentiate human genes from genes of other species, and are preferably based on unique nucleotide sequences.

The sequences falling within the scope of the present invention are not limited to these specific sequences, but also include allelic and species variations thereof. Allelic and species variations can be routinely determined by comparing the sequence provided in SEQ ID NO: 1 - 948, a representative fragment thereof, or a nucleotide sequence at least 90% identical, preferably 95% identical, to SEQ ID NOs: 1 - 948 with a sequence from another isolate of the same species. Furthermore, to accommodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another codon that encodes the same amino acid is expressly contemplated.

The nearest neighbor or homology result for the nucleic acids of the present invention, including SEQ ID NOs: 1 - 948, can be obtained by searching a database using an algorithm or a program. Preferably, a BLAST which stands for Basic Local Alignment Search Tool is used to search for local sequence alignments (Altshul, S.F. J Mol. Evol. 36 290-300 (1993) and Altschul S.F. et al. J. Mol. Biol. 21:403-410 (1990)). Alternatively a FASTA version 3 search against Genpept, using Fastxy algorithm.

Species homologs (or orthologs) of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides.

The nucleic acid sequences of the invention are further directed to sequences which encode variants of the described nucleic acids. These amino acid sequence variants may be prepared by methods known in the art by introducing appropriate nucleotide changes into a native or variant polynucleotide. There are two variables in the construction of amino acid sequence variants: the location of the mutation and the nature of the mutation. Nucleic acids encoding the amino acid sequence variants are preferably constructed by mutating the polynucleotide to encode an amino acid sequence that does not occur in nature. These nucleic acid alterations can be made at sites that differ in the nucleic acids from different species (variable positions) or in highly conserved regions (constant regions). Sites at such locations will typically be modified in series, *e.g.*, by substituting first with conservative choices (*e.g.*, hydrophobic amino acid to a different hydrophobic amino acid) and then with

more distant choices (*e.g.*, hydrophobic amino acid to a charged amino acid), and then deletions or insertions may be made at the target site. Amino acid sequence deletions generally range from about 1 to 30 residues, preferably about 1 to 10 residues, and are typically contiguous. Amino acid insertions include amino- and/or carboxyl-terminal fusions ranging in length from one to one hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions may range generally from about 1 to 10 amino residues, preferably from 1 to 5 residues. Examples of terminal insertions include the heterologous signal sequences necessary for secretion or for intracellular targeting in different host cells and sequences such as FLAG or poly-histidine sequences useful for purifying the expressed protein.

In a preferred method, polynucleotides encoding the novel amino acid sequences are changed via site-directed mutagenesis. This method uses oligonucleotide sequences to alter a polynucleotide to encode the desired amino acid variant, as well as sufficient adjacent nucleotides on both sides of the changed amino acid to form a stable duplex on either side of the site of being changed. In general, the techniques of site-directed mutagenesis are well known to those of skill in the art and this technique is exemplified by publications such as, Edelman et al., *DNA* 2:183 (1983). A versatile and efficient method for producing site-specific changes in a polynucleotide sequence was published by Zoller and Smith, *Nucleic Acids Res.* 10:6487-6500 (1982). PCR may also be used to create amino acid sequence variants of the novel nucleic acids. When small amounts of template DNA are used as starting material, primer(s) that differs slightly in sequence from the corresponding region in the template DNA can generate the desired amino acid variant. PCR amplification results in a population of product DNA fragments that differ from the polynucleotide template encoding the polypeptide at the position specified by the primer. The product DNA fragments replace the corresponding region in the plasmid and this gives a polynucleotide encoding the desired amino acid variant.

A further technique for generating amino acid variants is the cassette mutagenesis technique described in Wells et al., *Gene* 34:315 (1985); and other mutagenesis techniques well known in the art, such as, for example, the techniques in Sambrook et al., supra, and *Current Protocols in Molecular Biology*, Ausubel et al. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be used in the practice of the invention for the cloning and expression of these novel nucleic acids. Such DNA sequences include those which are

capable of hybridizing to the appropriate novel nucleic acid sequence under stringent conditions.

Polynucleotides encoding preferred polypeptide truncations of the invention can be used to generate polynucleotides encoding chimeric or fusion proteins comprising one or more domains of the invention and heterologous protein sequences.

The polynucleotides of the invention additionally include the complement of any of the polynucleotides recited above. The polynucleotide can be DNA (genomic, cDNA, amplified, or synthetic) or RNA. Methods and algorithms for obtaining such polynucleotides are well known to those of skill in the art and can include, for example, methods for determining hybridization conditions that can routinely isolate polynucleotides of the desired sequence identities.

In accordance with the invention, polynucleotide sequences comprising the mature protein coding sequences corresponding to any one of SEQ ID NO: 1-948, or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of that nucleic acid, or a functional equivalent thereof, in appropriate host cells. Also included are the cDNA inserts of any of the clones identified herein.

A polynucleotide according to the invention can be joined to any of a variety of other nucleotide sequences by well-established recombinant DNA techniques (see Sambrook J et al. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY). Useful nucleotide sequences for joining to polynucleotides include an assortment of vectors, e.g., plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a polynucleotide of the invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism, convenient restriction endonuclease sites, and a selectable marker for the host cell. Vectors according to the invention include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

The present invention further provides recombinant constructs comprising a nucleic acid having any of the nucleotide sequences of SEQ ID NOs: 1 - 948 or a fragment thereof or any other polynucleotides of the invention. In one embodiment, the recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a nucleic acid having any of the nucleotide sequences of SEQ ID NOs: 1 - 948 or a fragment thereof is inserted, in a forward or reverse orientation. In the case of a vector comprising one

of the ORFs of the present invention, the vector may further comprise regulatory sequences, including for example, a promoter, operably linked to the ORF. Large numbers of suitable vectors and promoters are known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of example. Bacterial: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLneo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al., *Nucleic Acids Res.* 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, *Methods in Enzymology* 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art. Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, *e.g.*, the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), a-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an amino terminal identification peptide imparting desired characteristics, *e.g.*, stabilization or simplified purification of expressed

recombinant product. Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced or derepressed by appropriate means (*e.g.*, temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Polynucleotides of the invention can also be used to induce immune responses. For example, as described in Fan et al., *Nat. Biotech.* 17:870-872 (1999), incorporated herein by reference, nucleic acid sequences encoding a polypeptide may be used to generate antibodies against the encoded polypeptide following topical administration of naked plasmid DNA or following injection, and preferably intra-muscular injection of the DNA. The nucleic acid sequences are preferably inserted in a recombinant expression vector and may be in the form of naked DNA.

3.3 ANTISENSE

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 1 - 948, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a

double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a protein of any of SEQ ID NO: 949-1896 or antisense nucleic acids complementary to a nucleic acid sequence of SEQ ID NO: 1 - 948 are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence of the invention. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence of the invention. The term "noncoding region" refers to 5' and 3' sequences that flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding a nucleic acid disclosed herein (*e.g.*, SEQ ID NO: 1 - 948, antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of an mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of an mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of an mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine,

inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a protein according to the invention to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual α-units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids Res* 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et*

al. (1987) *Nucleic Acids Res* 15: 6131-6148) or a chimeric RNA -DNA analogue (Inoue *et al.* (1987) *FEBS Lett* 215: 327-330).

3.4 RIBOZYMES AND PNA MOIETIES

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave mRNA transcripts to thereby inhibit translation of an mRNA. A ribozyme having specificity for a nucleic acid of the invention can be designed based upon the nucleotide sequence of a DNA disclosed herein (*i.e.*, SEQ ID NO: 1 - 948). For example, a derivative of Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a SECX-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Pat. No. 4,987,071; and Cech *et al.* U.S. Pat. No. 5,116,742. Alternatively, SECX mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel *et al.*, (1993) *Science* 261:1411-1418.

Alternatively, gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region (*e.g.*, promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells. See generally, Helene. (1991) *Anticancer Drug Des.* 6: 569-84; Helene. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14: 807-15.

In various embodiments, the nucleic acids of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup *et al.* (1996) *Bioorg Med Chem* 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup *et al.* (1996) above; Perry-O'Keefe *et al.* (1996) *PNAS* 93: 14670-675.

PNAs of the invention can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. PNAs of the invention can also be used, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup *et al.* (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of the invention can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated that may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, *e.g.*, RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn *et al.* (1996) *Nucl Acids Res* 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag *et al.* (1989) *Nucl Acid Res* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn *et al.* (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen *et al.* (1975) *Bioorg Med Chem Lett* 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.*, 1989, *Proc. Natl. Acad. Sci. U.S.A.* 86:6553-6556; Lemaitre *et al.*, 1987, Proc. *Natl. Acad. Sci.* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, *e.g.*, Krol *et al.*, 1988, *BioTechniques* 6:958-976) or intercalating agents. (See, *e.g.*, Zon, 1988, *Pharm. Res.* 5: 539-549). To this end, the oligonucleotide may be conjugated to

another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

3.5 HOSTS

The present invention further provides host cells genetically engineered to contain the polynucleotides of the invention. For example, such host cells may contain nucleic acids of the invention introduced into the host cell using known transformation, transfection or infection methods. The present invention still further provides host cells genetically engineered to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell.

Knowledge of nucleic acid sequences allows for modification of cells to permit, or increase, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the polypeptide at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the encoding sequences. See, for example, PCT International Publication No. WO94/12650, PCT International Publication No. WO92/20808, and PCT International Publication No. WO91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the recombinant construct into the host cell can be effected by calcium phosphate transfection, DEAE, dextran mediated transfection, or electroporation (Davis, L. et al., *Basic Methods in Molecular Biology* (1986)). The host cells containing one of the polynucleotides of the invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

Any host/vector system can be used to express one or more of the ORFs of the present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, Cv-1 cell, COS cells, 293 cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular polypeptide or protein or which expresses the polypeptide or protein at low natural level. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., in Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, New York (1989), the disclosure of which is hereby incorporated by reference.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell 23:175 (1981). Other cell lines capable of expressing a compatible vector are, for example, the C127, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5′ flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Recombinant polypeptides and proteins produced in bacterial culture are usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or insects or in prokaryotes such as bacteria. Potentially suitable yeast strains include

Saccharomyces cerevisiae, *Schizosaccharomyces pombe*, *Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli*, *Bacillus subtilis*, *Salmonella typhimurium*, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequence include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, *e.g.*, inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the host cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but

configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

3.6 POLYPEPTIDES OF THE INVENTION

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising: the amino acid sequences set forth as any one of SEQ ID NO: 949-1896 or an amino acid sequence encoded by any one of the nucleotide sequences SEQ ID NOs: 1 - 948 or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides preferably with biological or immunological activity that are encoded by: (a) a polynucleotide having any one of the nucleotide sequences set forth in SEQ ID NOs: 1 – 948 or (b) polynucleotides encoding any one of the amino acid sequences set forth as SEQ ID NO: 949-1896 or (c) polynucleotides that hybridize to the complement of the polynucleotides of either (a) or (b) under stringent hybridization conditions. The invention also provides biologically active or immunologically active variants of any of the amino acid sequences set forth as SEQ ID NO: 949-1896 or the corresponding full length or mature protein; and "substantial equivalents" thereof (e.g., with at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, 86%, 87%, 88%, 89%, at least about 90%, 91%, 92%, 93%, 94%, typically at least about 95%, 96%, 97%, more typically at least about 98%, or most typically at least about 99% amino acid identity) that retain biological activity. Polypeptides encoded by allelic variants may have a similar, increased, or decreased activity compared to polypeptides comprising SEQ ID NO: 949-1896.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H. U. Saragovi, et al., Bio/Technology 10, 773-778 (1992) and in R. S.

McDowell, et al., J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites.

The present invention also provides both full-length and mature forms (for example, without a signal sequence or precursor sequence) of the disclosed proteins. The protein coding sequence is identified in the sequence listing by translation of the disclosed nucleotide sequences. The mature form of such protein may be obtained by expression of a full-length polynucleotide in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein is also determinable from the amino acid sequence of the full-length form. Where proteins of the present invention are membrane bound, soluble forms of the proteins are also provided. In such forms, part or all of the regions causing the proteins to be membrane bound are deleted so that the proteins are fully secreted from the cell in which they are expressed.

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, *e.g.*, pharmaceutically acceptable, carrier.

The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. By "degenerate variant" is intended nucleotide fragments which differ from a nucleic acid fragment of the present invention (*e.g.*, an ORF) by nucleotide sequence but, due to the degeneracy of the genetic code, encode an identical polypeptide sequence. Preferred nucleic acid fragments of the present invention are the ORFs that encode proteins.

A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. This technique is particularly useful in producing small peptides and fragments of larger polypeptides. Fragments are useful, for example, in generating antibodies against the native polypeptide. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The polypeptides and proteins of the present invention can alternatively be purified from cells which have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. One skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic sequences into eukaryotic or prokaryotic cells in order to generate a cell which produces one of the polypeptides or proteins of the present invention.

The invention also relates to methods for producing a polypeptide comprising growing a culture of host cells of the invention in a suitable culture medium, and purifying the protein from the cells or the culture in which the cells are grown. For example, the methods of the invention include a process for producing a polypeptide in which a host cell containing a suitable expression vector that includes a polynucleotide of the invention is cultured under conditions that allow expression of the encoded polypeptide. The polypeptide can be recovered from the culture, conveniently from the culture medium, or from a lysate prepared from the host cells and further purified. Preferred embodiments include those in which the protein produced by such process is a full length or mature form of the protein.

In an alternative method, the polypeptide or protein is purified from bacterial cells which naturally produce the polypeptide or protein. One skilled in the art can readily follow known methods for isolating polypeptides and proteins in order to obtain one of the isolated polypeptides or proteins of the present invention. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immuno-affinity chromatography. See, *e.g.*, Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag (1994); Sambrook, et al., in Molecular Cloning: *A Laboratory Manual*; Ausubel et al., *Current Protocols in Molecular Biology*. Polypeptide fragments that retain biological/immunological activity include fragments comprising greater than about 100 amino acids, or greater than about 200 amino acids, and fragments that encode specific protein domains.

The purified polypeptides can be used in *in vitro* binding assays which are well known in the art to identify molecules which bind to the polypeptides. These molecules include but are not limited to, for e.g., small molecules, molecules from combinatorial libraries, antibodies or other proteins. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well

known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

In addition, the peptides of the invention or molecules capable of binding to the peptides may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for SEQ ID NO: 949-1896.

The protein of the invention may also be expressed as a product of transgenic animals, *e.g.*, as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications, in the peptide or DNA sequence, can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, *e.g.*, U.S. Pat. No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein. Regions of the protein that are important for the protein function can be determined by various methods known in the art including the alanine-scanning method which involved systematic substitution of single or strings of amino acids with alanine, followed by testing the resulting alanine-containing variant for biological activity. This type of analysis determines the importance of the substituted amino acid(s) in biological activity. Regions of the protein that are important for protein function may be determined by the eMATRIX program.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and are useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are encompassed by the present invention.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *e.g.*, Invitrogen, San Diego,

Calif., U.S.A. (the MaxBat™ kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (*i.e.*, from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl™ or Cibacrom blue 3GA Sepharose™; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX), or as a His tag. Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, Mass.), Pharmacia (Piscataway, N.J.) and Invitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("FLAG®") is commercially available from Kodak (New Haven, Conn.).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, *e.g.*, silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The polypeptides of the invention include analogs (variants). This embraces fragments, as well as peptides in which one or more amino acids has been deleted, inserted, or substituted. Also, analogs of the polypeptides of the invention embrace fusions of the polypeptides or modifications of the polypeptides of the invention, wherein the polypeptide or analog is fused to another moiety or moieties, e.g., targeting moiety or another therapeutic agent. Such analogs may exhibit improved properties such as activity and/or stability.

Examples of moieties which may be fused to the polypeptide or an analog include, for example, targeting moieties which provide for the delivery of polypeptide to pancreatic cells, e.g., antibodies to pancreatic cells, antibodies to immune cells such as T-cells, monocytes, dendritic cells, granulocytes, etc., as well as receptor and ligands expressed on pancreatic or immune cells. Other moieties which may be fused to the polypeptide include therapeutic agents which are used for treatment, for example, immunosuppressive drugs such as cyclosporin, SK506, azathioprine, CD3 antibodies and steroids. Also, polypeptides may be fused to immune modulators, and other cytokines such as alpha or beta interferon.

3.6.1 DETERMINING POLYPEPTIDE AND POLYNUCLEOTIDE IDENTITY AND SIMILARITY

Preferred identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in computer programs including, but are not limited to, the GCG program package, including GAP (Devereux, J., et al., Nucleic Acids Research 12(1):387 (1984); Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, BLASTX, FASTA (Altschul, S.F. et al., J. Molec. Biol. 215:403-410 (1990), PSI-BLAST (Altschul S.F. et al., Nucleic Acids Res. vol. 25, pp. 3389-3402, herein incorporated by reference), eMatrix software (Wu et al., J. Comp. Biol., Vol. 6, pp. 219-235 (1999), herein incorporated by reference), eMotif software (Nevill-Manning et al, ISMB-97, Vol. 4, pp. 202-209, herein incorporated by reference), pFam software (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1), pp. 320-322 (1998), herein incorporated by reference) and the Kyte-Doolittle hydrophobocity prediction algorithm (J. Mol Biol, 157, pp. 105-31 (1982), incorporated herein by reference). The BLAST programs are publicly available from the National Center for Biotechnology Information (NCBI) and other sources (BLAST Manual, Altschul, S., et al. NCB NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990).

3.7 CHIMERIC AND FUSION PROTEINS

The invention also provides chimeric or fusion proteins. As used herein, a "chimeric protein" or "fusion protein" comprises a polypeptide of the invention operatively linked to another polypeptide. Within a fusion protein the polypeptide according to the invention can correspond to all or a portion of a protein according to the invention. In one embodiment, a fusion protein comprises at least one biologically active portion of a protein according to the invention. In another embodiment, a fusion protein comprises at least two biologically active

portions of a protein according to the invention. Within the fusion protein, the term "operatively linked" is intended to indicate that the polypeptide according to the invention and the other polypeptide are fused in-frame to each other. The polypeptide can be fused to the N-terminus or C-terminus, or to the middle.

For example, in one embodiment a fusion protein comprises a polypeptide according to the invention operably linked to the extracellular domain of a second protein.

In another embodiment, the fusion protein is a GST-fusion protein in which the polypeptide sequences of the invention are fused to the C-terminus of the GST (i.e., glutathione S-transferase) sequences.

In another embodiment, the fusion protein is an immunoglobulin fusion protein in which the polypeptide sequences according to the invention comprise one or more domains fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ligand and a protein of the invention on the surface of a cell, to thereby suppress signal transduction *in vivo*. The immunoglobulin fusion proteins can be used to affect the bioavailability of a cognate ligand. Inhibition of the ligand/protein interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, *e.g.*, cancer as well as modulating (*e.g.*, promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies in a subject, to purify ligands, and in screening assays to identify molecules that inhibit the interaction of a polypeptide of the invention with a ligand.

A chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, *e.g.*, by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, Ausubel et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many

expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A nucleic acid encoding a polypeptide of the invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the protein of the invention.

3.8 GENE THERAPY

Mutations in the polynucleotides of the invention gene may result in loss of normal function of the encoded protein. The invention thus provides gene therapy to restore normal activity of the polypeptides of the invention; or to treat disease states involving polypeptides of the invention. Delivery of a functional gene encoding polypeptides of the invention to appropriate cells is effected *ex vivo*, *in situ*, or *in vivo* by use of vectors, and more particularly viral vectors (e.g., adenovirus, adeno-associated virus, or a retrovirus), or *ex vivo* by use of physical DNA transfer methods (e.g., liposomes or chemical treatments). See, for example, Anderson, Nature, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, Science, 244: 1275-1281 (1989); Verma, Scientific American: 68-84 (1990); and Miller, Nature, 357: 455-460 (1992). Introduction of any one of the nucleotides of the present invention or a gene encoding the polypeptides of the present invention can also be accomplished with extrachromosomal substrates (transient expression) or artificial chromosomes (stable expression). Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes. Alternatively, it is contemplated that in other human disease states, preventing the expression of or inhibiting the activity of polypeptides of the invention will be useful in treating the disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of polypeptides of the invention.

Other methods inhibiting expression of a protein include the introduction of antisense molecules to the nucleic acids of the present invention, their complements, or their translated RNA sequences, by methods known in the art. Further, the polypeptides of the present invention can be inhibited by using targeted deletion methods, or the insertion of a negative regulatory element such as a silencer, which is tissue specific.

The present invention still further provides cells genetically engineered *in vivo* to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the

polynucleotides in the cell. These methods can be used to increase or decrease the expression of the polynucleotides of the present invention.

Knowledge of DNA sequences provided by the invention allows for modification of cells to permit, increase, or decrease, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the protein at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the desired protein encoding sequences. See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., *ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase*, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the desired protein coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequences include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, *e.g.*, inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element.

Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

3.9 TRANSGENIC ANIMALS

In preferred methods to determine biological functions of the polypeptides of the invention in vivo, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals,

preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of a promoter of the polynucleotides of the invention is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

The polynucleotides of the present invention also make possible the development, through, e.g., homologous recombination or knock out strategies, of animals that fail to express polypeptides of the invention or that express a variant polypeptide. Such animals are useful as models for studying the *in vivo* activities of polypeptide as well as for studying modulators of the polypeptides of the invention.

In preferred methods to determine biological functions of the polypeptides of the invention *in vivo*, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of the polynucleotides of the invention promoter is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

3.10 USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified herein. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA). The mechanism underlying the particular condition or pathology will dictate whether the polypeptides of the invention, the polynucleotides of the invention or modulators (activators or inhibitors) thereof would be beneficial to the subject in need of treatment. Thus, "therapeutic compositions of the invention" include compositions comprising isolated polynucleotides (including recombinant DNA molecules, cloned genes and degenerate variants thereof) or polypeptides of the invention (including full length protein, mature protein and truncations or domains thereof), or compounds and other substances that modulate the overall activity of the target gene products, either at the level of target gene/protein expression or target protein activity. Such modulators include polypeptides, analogs, (variants), including fragments and fusion proteins, antibodies and other binding proteins; chemical compounds that directly or indirectly activate or inhibit the polypeptides of the invention (identified, e.g., via drug screening assays as described herein); antisense polynucleotides and polynucleotides suitable for triple helix formation; and in particular antibodies or other binding partners that specifically recognize one or more epitopes of the polypeptides of the invention.

The polypeptides of the present invention may likewise be involved in cellular activation or in one of the other physiological pathways described herein.

3.10.1 RESEARCH USES AND UTILITIES

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA

sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The polypeptides provided by the present invention can similarly be used in assays to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding polypeptide is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

3.10.2 NUTRITIONAL USES

Polynucleotides and polypeptides of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the polypeptide or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the

polypeptide or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

3.10.3 CYTOKINE AND CELL PROLIFERATION/DIFFERENTIATION ACTIVITY

A polypeptide of the present invention may exhibit activity relating to cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor-dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of therapeutic compositions of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+(preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e, CMK, HUVEC, and Caco. Therapeutic compositions of the invention can be used in the following:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., I. Immunol. 149:3778-3783, 1992; Bowman et al., I. Immunol. 152:1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A. M. and Shevach, E. M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human interleukin-γ, Schreiber, R. D. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L. S. and Lipsky, P. E. In Current

Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6--Nordan, R. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Aced. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11--Bennett, F., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9--Ciarletta, A., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

3.10.4 STEM CELL GROWTH FACTOR ACTIVITY

A polypeptide of the present invention may exhibit stem cell growth factor activity and be involved in the proliferation, differentiation and survival of pluripotent and totipotent stem cells including primordial germ cells, embryonic stem cells, hematopoietic stem cells and/or germ line stem cells. Administration of the polypeptide of the invention to stem cells *in vivo* or *ex vivo* is expected to maintain and expand cell populations in a totipotential or pluripotential state which would be useful for re-engineering damaged or diseased tissues, transplantation, manufacture of bio-pharmaceuticals and the development of bio-sensors. The ability to produce large quantities of human cells has important working applications for the production of human proteins which currently must be obtained from non-human sources or donors, implantation of cells to treat diseases such as Parkinson's, Alzheimer's and other neurodegenerative diseases; tissues for grafting such as bone marrow, skin, cartilage, tendons, bone, muscle (including cardiac muscle), blood vessels, cornea, neural cells, gastrointestinal

cells and others; and organs for transplantation such as kidney, liver, pancreas (including islet cells), heart and lung.

It is contemplated that multiple different exogenous growth factors and/or cytokines may be administered in combination with the polypeptide of the invention to achieve the desired effect, including any of the growth factors listed herein, other stem cell maintenance factors, and specifically including stem cell factor (SCF), leukemia inhibitory factor (LIF), Flt-3 ligand (Flt-3L), any of the interleukins, recombinant soluble IL-6 receptor fused to IL-6, macrophage inflammatory protein 1-alpha (MIP-1-alpha), G-CSF, GM-CSF, thrombopoietin (TPO), platelet factor 4 (PF-4), platelet-derived growth factor (PDGF), neural growth factors and basic fibroblast growth factor (bFGF).

Since totipotent stem cells can give rise to virtually any mature cell type, expansion of these cells in culture will facilitate the production of large quantities of mature cells. Techniques for culturing stem cells are known in the art and administration of polypeptides of the invention, optionally with other growth factors and/or cytokines, is expected to enhance the survival and proliferation of the stem cell populations. This can be accomplished by direct administration of the polypeptide of the invention to the culture medium. Alternatively, stroma cells transfected with a polynucleotide that encodes for the polypeptide of the invention can be used as a feeder layer for the stem cell populations in culture or in vivo. Stromal support cells for feeder layers may include embryonic bone marrow fibroblasts, bone marrow stromal cells, fetal liver cells, or cultured embryonic fibroblasts (see U.S. Patent No. 5,690,926).

Stem cells themselves can be transfected with a polynucleotide of the invention to induce autocrine expression of the polypeptide of the invention. This will allow for generation of undifferentiated totipotential/pluripotential stem cell lines that are useful as is or that can then be differentiated into the desired mature cell types. These stable cell lines can also serve as a source of undifferentiated totipotential/pluripotential mRNA to create cDNA libraries and templates for polymerase chain reaction experiments. These studies would allow for the isolation and identification of differentially expressed genes in stem cell populations that regulate stem cell proliferation and/or maintenance.

Expansion and maintenance of totipotent stem cell populations will be useful in the treatment of many pathological conditions. For example, polypeptides of the present invention may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease, accidental damage or genetic disorders. The polypeptide of the invention may be useful for

inducing the proliferation of neural cells and for the regeneration of nerve and brain tissue, i.e. for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders which involve degeneration, death or trauma to neural cells or nerve tissue. In addition, the expanded stem cell populations can also be genetically altered for gene therapy purposes and to decrease host rejection of replacement tissues after grafting or implantation.

Expression of the polypeptide of the invention and its effect on stem cells can also be manipulated to achieve controlled differentiation of the stem cells into more differentiated cell types. A broadly applicable method of obtaining pure populations of a specific differentiated cell type from undifferentiated stem cell populations involves the use of a cell-type specific promoter driving a selectable marker. The selectable marker allows only cells of the desired type to survive. For example, stem cells can be induced to differentiate into cardiomyocytes (Wobus et al., Differentiation, 48: 173-182, (1991); Klug et al., J. Clin. Invest., 98(1): 216-224, (1998)) or skeletal muscle cells (Browder, L. W. In: *Principles of Tissue Engineering eds.* Lanza et al., Academic Press (1997)). Alternatively, directed differentiation of stem cells can be accomplished by culturing the stem cells in the presence of a differentiation factor such as retinoic acid and an antagonist of the polypeptide of the invention which would inhibit the effects of endogenous stem cell factor activity and allow differentiation to proceed.

In vitro cultures of stem cells can be used to determine if the polypeptide of the invention exhibits stem cell growth factor activity. Stem cells are isolated from any one of various cell sources (including hematopoietic stem cells and embryonic stem cells) and cultured on a feeder layer, as described by Thompson et al. Proc. Natl. Acad. Sci, U.S.A., 92: 7844-7848 (1995), in the presence of the polypeptide of the invention alone or in combination with other growth factors or cytokines. The ability of the polypeptide of the invention to induce stem cells proliferation is determined by colony formation on semi-solid support e.g. as described by Bernstein et al., Blood, 77: 2316-2321 (1991).

3.10.5 HEMATOPOIESIS REGULATING ACTIVITY

A polypeptide of the present invention may be involved in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell disorders. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines,

thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

Therapeutic compositions of the invention can be used in the following:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M. G. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I. K. and Briddell, R. A. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R. E. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Long term bone marrow cultures

in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Long term culture initiating cell assay, Sutherland, H. J. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

3.10.6 TISSUE GROWTH ACTIVITY

A polypeptide of the present invention also may be involved in bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as in wound healing and tissue repair and replacement, and in healing of burns, incisions and ulcers.

A polypeptide of the present invention which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Compositions of a polypeptide, antibody, binding partner, or other modulator of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A polypeptide of this invention may also be involved in attracting bone-forming cells, stimulating growth of bone-forming cells, or inducing differentiation of progenitors of bone-forming cells. Treatment of osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes may also be possible using the composition of the invention.

Another category of tissue regeneration activity that may involve the polypeptide of the present invention is tendon/ligament formation. Induction of tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention

contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The compositions of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, i.e. for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a composition may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a composition of the invention.

Compositions of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

Compositions of the present invention may also be involved in the generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring may allow normal tissue to regenerate. A polypeptide of the present invention may also exhibit angiogenic activity.

A composition of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A composition of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

Therapeutic compositions of the invention can be used in the following:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, H. I. and Rovee, D. T., eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

3.10.7 IMMUNE STIMULATING OR SUPPRESSING ACTIVITY

A polypeptide of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A polynucleotide of the invention can encode a polypeptide exhibiting such activities. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpes viruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, proteins of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein (or antagonists thereof, including antibodies) of the present invention may also to be useful in the

treatment of allergic reactions and conditions (*e.g.*, anaphylaxis, serum sickness, drug reactions, food allergies, insect venom allergies, mastocytosis, allergic rhinitis, hypersensitivity pneumonitis, urticaria, angioedema, eczema, atopic dermatitis, allergic contact dermatitis, erythema multiforme, Stevens-Johnson syndrome, allergic conjunctivitis, atopic keratoconjunctivitis, venereal keratoconjunctivitis, giant papillary conjunctivitis and contact allergies), such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein (or antagonists thereof) of the present invention. The therapeutic effects of the polypeptides or antagonists thereof on allergic reactions can be evaluated by in vivo animals models such as the cumulative contact enhancement test (Lastbom et al., Toxicology 125: 59-66, 1998), skin prick test (Hoffmann et al., Allergy 54: 446-54, 1999), guinea pig skin sensitization test (Vohr et al., Arch. Toxocol. 73: 501-9), and murine local lymph node assay (Kimber et al., J. Toxicol. Environ. Health 53: 563-79).

Using the proteins of the invention it may also be possible to modulate immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as, for example, B7)), e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a therapeutic composition of the invention may prevent cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, a lack of costimulation

may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular therapeutic compositions in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins in vivo as described in Lenschow et al., Science 257:789-792 (1992) and Turka et al., Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of therapeutic compositions of the invention on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block stimulation of T cells can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (e.g., a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response may be useful in

cases of viral infection, including systemic viral diseases such as influenza, the common cold, and encephalitis.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells in vitro with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the in vitro activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells in vivo.

A polypeptide of the present invention may provide the necessary stimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient mounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I alpha chain protein and β_2 microglobulin protein or an MHC class II alpha chain protein and an MHC class II beta chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19;

Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., I. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bowman et al., J. Virology 61:1992-1998; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: In vitro antibody production, Mond, J. J. and Brunswick, M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of

Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Gaily et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

3.10.8 ACTIVIN/INHIBIN ACTIVITY

A polypeptide of the present invention may also exhibit activin- or inhibin-related activities. A polynucleotide of the invention may encode a polypeptide exhibiting such characteristics. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a polypeptide of the present invention, alone or in heterodimers with a member of the inhibin family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the polypeptide of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A polypeptide of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as, but not limited to, cows, sheep and pigs.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods.

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

3.10.9 CHEMOTACTIC/CHEMOKINETIC ACTIVITY

A polypeptide of the present invention may be involved in chemotactic or chemokinetic activity for mammalian cells, including, for example, monocytes, fibroblasts,

neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Chemotactic and chemokinetic receptor activation can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic compositions (e.g. proteins, antibodies, binding partners, or modulators of the invention) provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

Therapeutic compositions of the invention can be used in the following:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Marguiles, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25:1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153:1762-1768, 1994.

3.10.10 HEMOSTATIC AND THROMBOLYTIC ACTIVITY

A polypeptide of the invention may also be involved in hemostatis or thrombolysis or thrombosis. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Compositions may be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A composition of the invention may also be useful for dissolving or inhibiting formation of

thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

Therapeutic compositions of the invention can be used in the following:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

3.10.11 CANCER DIAGNOSIS AND THERAPY

Polypeptides of the invention may be involved in cancer cell generation, proliferation or metastasis. Detection of the presence or amount of polynucleotides or polypeptides of the invention may be useful for the diagnosis and/or prognosis of one or more types of cancer. For example, the presence or increased expression of a polynucleotide/polypeptide of the invention may indicate a hereditary risk of cancer, a precancerous condition, or an ongoing malignancy. Conversely, a defect in the gene or absence of the polypeptide may be associated with a cancer condition. Identification of single nucleotide polymorphisms associated with cancer or a predisposition to cancer may also be useful for diagnosis or prognosis.

Cancer treatments promote tumor regression by inhibiting tumor cell proliferation, inhibiting angiogenesis (growth of new blood vessels that is necessary to support tumor growth) and/or prohibiting metastasis by reducing tumor cell motility or invasiveness. Therapeutic compositions of the invention may be effective in adult and pediatric oncology including in solid phase tumors/malignancies, locally advanced tumors, human soft tissue sarcomas, metastatic cancer, including lymphatic metastases, blood cell malignancies including multiple myeloma, acute and chronic leukemias, and lymphomas, head and neck cancers including mouth cancer, larynx cancer and thyroid cancer, lung cancers including small cell carcinoma and non-small cell cancers, breast cancers including small cell carcinoma and ductal carcinoma, gastrointestinal cancers including esophageal cancer, stomach cancer, colon cancer, colorectal cancer and polyps associated with colorectal neoplasia, pancreatic cancers, liver cancer, urologic cancers including bladder cancer and prostate cancer, malignancies of the female genital tract including ovarian carcinoma, uterine (including endometrial) cancers, and solid tumor in the ovarian follicle, kidney cancers including renal cell carcinoma, brain cancers including intrinsic brain tumors, neuroblastoma, astrocytic brain tumors, gliomas, metastatic tumor cell invasion in the central nervous system,

bone cancers including osteomas, skin cancers including malignant melanoma, tumor progression of human skin keratinocytes, squamous cell carcinoma, basal cell carcinoma, hemangiopericytoma and Karposi's sarcoma.

Polypeptides, polynucleotides, or modulators of polypeptides of the invention (including inhibitors and stimulators of the biological activity of the polypeptide of the invention) may be administered to treat cancer. Therapeutic compositions can be administered in therapeutically effective dosages alone or in combination with adjuvant cancer therapy such as surgery, chemotherapy, radiotherapy, thermotherapy, and laser therapy, and may provide a beneficial effect, e.g. reducing tumor size, slowing rate of tumor growth, inhibiting metastasis, or otherwise improving overall clinical condition, without necessarily eradicating the cancer.

The composition can also be administered in therapeutically effective amounts as a portion of an anti-cancer cocktail. An anti-cancer cocktail is a mixture of the polypeptide or modulator of the invention with one or more anti-cancer drugs in addition to a pharmaceutically acceptable carrier for delivery. The use of anti-cancer cocktails as a cancer treatment is routine. Anti-cancer drugs that are well known in the art and can be used as a treatment in combination with the polypeptide or modulator of the invention include: Actinomycin D, Aminoglutethimide, Asparaginase, Bleomycin, Busulfan, Carboplatin, Carmustine, Chlorambucil, Cisplatin (cis-DDP), Cyclophosphamide, Cytarabine HCl (Cytosine arabinoside), Dacarbazine, Dactinomycin, Daunorubicin HCl, Doxorubicin HCl, Estramustine phosphate sodium, Etoposide (V16-213), Floxuridine, 5-Fluorouracil (5-Fu), Flutamide, Hydroxyurea (hydroxycarbamide), Ifosfamide, Interferon Alpha-2a, Interferon Alpha-2b, Leuprolide acetate (LHRH-releasing factor analog), Lomustine, Mechlorethamine HCl (nitrogen mustard), Melphalan, Mercaptopurine, Mesna, Methotrexate (MTX), Mitomycin, Mitoxantrone HCl, Octreotide, Plicamycin, Procarbazine HCl, Streptozocin, Tamoxifen citrate, Thioguanine, Thiotepa, Vinblastine sulfate, Vincristine sulfate, Amsacrine, Azacitidine, Hexamethylmelamine, Interleukin-2, Mitoguazone, Pentostatin, Semustine, Teniposide, and Vindesine sulfate.

In addition, therapeutic compositions of the invention may be used for prophylactic treatment of cancer. There are hereditary conditions and/or environmental situations (e.g. exposure to carcinogens) known in the art that predispose an individual to developing cancers. Under these circumstances, it may be beneficial to treat these individuals with therapeutically effective doses of the polypeptide of the invention to reduce the risk of developing cancers.

In vitro models can be used to determine the effective doses of the polypeptide of the invention as a potential cancer treatment. These *in vitro* models include proliferation assays of cultured tumor cells, growth of cultured tumor cells in soft agar (see Freshney, (1987) Culture of Animal Cells: A Manual of Basic Technique, Wily-Liss, New York, NY Ch 18 and Ch 21), tumor systems in nude mice as described in Giovanella et al., J. Natl. Can. Inst., 52: 921-30 (1974), mobility and invasive potential of tumor cells in Boyden Chamber assays as described in Pilkington et al., Anticancer Res., 17: 4107-9 (1997), and angiogenesis assays such as induction of vascularization of the chick chorioallantoic membrane or induction of vascular endothelial cell migration as described in Ribatta et al., Intl. J. Dev. Biol., 40: 1189-97 (1999) and Li et al., Clin. Exp. Metastasis, 17:423-9 (1999), respectively. Suitable tumor cells lines are available, e.g. from American Type Tissue Culture Collection catalogs.

3.10.12 RECEPTOR/LIGAND ACTIVITY

A polypeptide of the present invention may also demonstrate activity as receptor, receptor ligand or inhibitor or agonist of receptor/ligand interactions. A polynucleotide of the invention can encode a polypeptide exhibiting such characteristics. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses. Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

By way of example, the polypeptides of the invention may be used as a receptor for a ligand(s) thereby transmitting the biological activity of that ligand(s). Ligands may be identified through binding assays, affinity chromatography, dihybrid screening assays, BIAcore assays, gel overlay assays, or other methods known in the art.

Studies characterizing drugs or proteins as agonist or antagonist or partial agonists or a partial antagonist require the use of other proteins as competing ligands. The polypeptides of the present invention or ligand(s) thereof may be labeled by being coupled to radioisotopes, colorimetric molecules or a toxin molecules by conventional methods. ("Guide to Protein Purification" Murray P. Deutscher (ed) Methods in Enzymology Vol. 182 (1990) Academic Press, Inc. San Diego). Examples of radioisotopes include, but are not limited to, tritium and carbon-14 . Examples of colorimetric molecules include, but are not limited to, fluorescent molecules such as fluorescamine, or rhodamine or other colorimetric molecules. Examples of toxins include, but are not limited, to ricin.

3.10.13 DRUG SCREENING

This invention is particularly useful for screening chemical compounds by using the novel polypeptides or binding fragments thereof in any of a variety of drug screening techniques. The polypeptides or fragments employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the polypeptide or a fragment thereof. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between polypeptides of the invention or fragments and the agent being tested or examine the diminution in complex formation between the novel polypeptides and an appropriate cell line, which are well known in the art.

Sources for test compounds that may be screened for ability to bind to or modulate (i.e., increase or decrease) the activity of polypeptides of the invention include (1) inorganic and organic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of either random or mimetic peptides, oligonucleotides or organic molecules.

Chemical libraries may be readily synthesized or purchased from a number of commercial sources, and may include structural analogs of known compounds or compounds that are identified as "hits" or "leads" via natural product screening.

The sources of natural product libraries are microorganisms (including bacteria and fungi), animals, plants or other vegetation, or marine organisms, and libraries of mixtures for screening may be created by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of the organisms themselves. Natural product libraries include polyketides, non-ribosomal peptides, and (non-naturally occurring) variants thereof. For a review, see *Science 282*:63-68 (1998).

Combinatorial libraries are composed of large numbers of peptides, oligonucleotides or organic compounds and can be readily prepared by traditional automated synthesis methods, PCR, cloning or proprietary synthetic methods. Of particular interest are peptide and oligonucleotide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, *Curr. Opin. Biotechnol.* 8:701-707 (1997). For reviews and examples of peptidomimetic libraries, see Al-Obeidi et al., *Mol. Biotechnol,* 9(3):205-23 (1998); Hruby et al., *Curr Opin Chem Biol,* 1(1):114-19 (1997); Dorner et al., *Bioorg Med Chem,* 4(5):709-15 (1996) (alkylated dipeptides).

Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to bind a polypeptide of the invention. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

The binding molecules thus identified may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells such as radioisotopes. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for a polypeptide of the invention. Alternatively, the binding molecules may be complexed with imaging agents for targeting and imaging purposes.

3.10.14 ASSAY FOR RECEPTOR ACTIVITY

The invention also provides methods to detect specific binding of a polypeptide e.g. a ligand or a receptor. The art provides numerous assays particularly useful for identifying previously unknown binding partners for receptor polypeptides of the invention. For example, expression cloning using mammalian or bacterial cells, or dihybrid screening assays can be used to identify polynucleotides encoding binding partners. As another example,

affinity chromatography with the appropriate immobilized polypeptide of the invention can be used to isolate polypeptides that recognize and bind polypeptides of the invention. There are a number of different libraries used for the identification of compounds, and in particular small molecules, that modulate (*i.e.,* increase or decrease) biological activity of a polypeptide of the invention. Ligands for receptor polypeptides of the invention can also be identified by adding exogenous ligands, or cocktails of ligands to two cells populations that are genetically identical except for the expression of the receptor of the invention: one cell population expresses the receptor of the invention whereas the other does not. The response of the two cell populations to the addition of ligands(s) are then compared. Alternatively, an expression library can be co-expressed with the polypeptide of the invention in cells and assayed for an autocrine response to identify potential ligand(s). As still another example, BIAcore assays, gel overlay assays, or other methods known in the art can be used to identify binding partner polypeptides, including, (1) organic and inorganic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules.

The role of downstream intracellular signaling molecules in the signaling cascade of the polypeptide of the invention can be determined. For example, a chimeric protein in which the cytoplasmic domain of the polypeptide of the invention is fused to the extracellular portion of a protein, whose ligand has been identified, is produced in a host cell. The cell is then incubated with the ligand specific for the extracellular portion of the chimeric protein, thereby activating the chimeric receptor. Known downstream proteins involved in intracellular signaling can then be assayed for expected modifications i.e. phosphorylation. Other methods known to those in the art can also be used to identify signaling molecules involved in receptor activity.

3.10.15 ANTI-INFLAMMATORY ACTIVITY

Compositions of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Compositions with such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation intimation

associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Compositions of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material. Compositions of this invention may be utilized to prevent or treat conditions such as, but not limited to, sepsis, acute pancreatitis, endotoxin shock, cytokine induced shock, rheumatoid arthritis, chronic inflammatory arthritis, pancreatic cell damage from diabetes mellitus type 1, graft versus host disease, inflammatory bowel disease, inflamation associated with pulmonary disease, other autoimmune disease or inflammatory disease, an antiproliferative agent such as for acute or chronic mylegenous leukemia or in the prevention of premature labor secondary to intrauterine infections.

3.10.16 LEUKEMIAS

Leukemias and related disorders may be treated or prevented by administration of a therapeutic that promotes or inhibits function of the polynucleotides and/or polypeptides of the invention. Such leukemias and related disorders include but are not limited to acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia, myeloblastic, promyelocytic, myelomonocytic, monocytic, erythroleukemia, chronic leukemia, chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia (for a review of such disorders, see Fishman et al., 1985, Medicine, 2d Ed., J.B. Lippincott Co., Philadelphia).

3.10.17 NERVOUS SYSTEM DISORDERS

Nervous system disorders, involving cell types which can be tested for efficacy of intervention with compounds that modulate the activity of the polynucleotides and/or polypeptides of the invention, and which can be treated upon thus observing an indication of therapeutic utility, include but are not limited to nervous system injuries, and diseases or disorders which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be treated in a patient (including human and non-human mammalian patients) according to the invention include but are not limited to the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems:

(i) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions which sever a portion of the nervous system, or compression injuries;

(ii) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or spinal cord infarction or ischemia;

(iii) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis;

(iv) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative process including but not limited to degeneration associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis;

(v) lesions associated with nutritional diseases or disorders, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration;

(vi) neurological lesions associated with systemic diseases including but not limited to diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis;

(vii) lesions caused by toxic substances including alcohol, lead, or particular neurotoxins; and

(viii) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including but not limited to multiple sclerosis, human immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

Therapeutics which are useful according to the invention for treatment of a nervous system disorder may be selected by testing for biological activity in promoting the survival or differentiation of neurons. For example, and not by way of limitation, therapeutics which elicit any of the following effects may be useful according to the invention:

(i) increased survival time of neurons in culture;

(ii) increased sprouting of neurons in culture or in vivo;

(iii) increased production of a neuron-associated molecule in culture or *in vivo*, *e.g.*, choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or

(iv) decreased symptoms of neuron dysfunction in vivo.

Such effects may be measured by any method known in the art. In preferred, non-limiting embodiments, increased survival of neurons may be measured by the method set forth in Arakawa et al. (1990, J. Neurosci. 10:3507-3515); increased sprouting of neurons may be detected by methods set forth in Pestronk et al. (1980, Exp. Neurol. 70:65-82) or Brown et al. (1981, Ann. Rev. Neurosci. 4:17-42); increased production of neuron-associated molecules may be measured by bioassay, enzymatic assay, antibody binding, Northern blot assay, *etc.*, depending on the molecule to be measured; and motor neuron dysfunction may be measured by assessing the physical manifestation of motor neuron disorder, *e.g.*, weakness, motor neuron conduction velocity, or functional disability.

In specific embodiments, motor neuron disorders that may be treated according to the invention include but are not limited to disorders such as infarction, infection, exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may affect motor neurons as well as other components of the nervous system, as well as disorders that selectively affect neurons such as amyotrophic lateral sclerosis, and including but not limited to progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, infantile and juvenile muscular atrophy, progressive bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio syndrome, and Hereditary Motorsensory Neuropathy (Charcot-Marie-Tooth Disease).

3.10.18 OTHER ACTIVITIES

A polypeptide of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, co-factors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition

(including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

3.10.19 IDENTIFICATION OF POLYMORPHISMS

The demonstration of polymorphisms makes possible the identification of such polymorphisms in human subjects and the pharmacogenetic use of this information for diagnosis and treatment. Such polymorphisms may be associated with, e.g., differential predisposition or susceptibility to various disease states (such as disorders involving inflammation or immune response) or a differential response to drug administration, and this genetic information can be used to tailor preventive or therapeutic treatment appropriately. For example, the existence of a polymorphism associated with a predisposition to inflammation or autoimmune disease makes possible the diagnosis of this condition in humans by identifying the presence of the polymorphism.

Polymorphisms can be identified in a variety of ways known in the art which all generally involve obtaining a sample from a patient, analyzing DNA from the sample, optionally involving isolation or amplification of the DNA, and identifying the presence of the polymorphism in the DNA. For example, PCR may be used to amplify an appropriate fragment of genomic DNA which may then be sequenced. Alternatively, the DNA may be subjected to allele-specific oligonucleotide hybridization (in which appropriate oligonucleotides are hybridized to the DNA under conditions permitting detection of a single base mismatch) or to a single nucleotide extension assay (in which an oligonucleotide that hybridizes immediately adjacent to the position of the polymorphism is extended with one or more labeled nucleotides). In addition, traditional restriction fragment length polymorphism analysis (using restriction enzymes that provide differential digestion of the genomic DNA depending on the presence or absence of the polymorphism) may be performed. Arrays with nucleotide sequences of the present invention can be used to detect polymorphisms. The array can comprise modified nucleotide sequences of the present invention in order to detect

the nucleotide sequences of the present invention. In the alternative, any one of the nucleotide sequences of the present invention can be placed on the array to detect changes from those sequences.

Alternatively a polymorphism resulting in a change in the amino acid sequence could also be detected by detecting a corresponding change in amino acid sequence of the protein, e.g., by an antibody specific to the variant sequence.

3.10.20 ARTHRITIS AND INFLAMMATION

The immunosuppressive effects of the compositions of the invention against rheumatoid arthritis is determined in an experimental animal model system. The experimental model system is adjuvant induced arthritis in rats, and the protocol is described by J. Holoshitz, et at., 1983, Science, 219:56, or by B. Waksman et al., 1963, Int. Arch. Allergy Appl. Immunol., 23:129. Induction of the disease can be caused by a single injection, generally intradermally, of a suspension of killed Mycobacterium tuberculosis in complete Freund's adjuvant (CFA). The route of injection can vary, but rats may be injected at the base of the tail with an adjuvant mixture. The polypeptide is administered in phosphate buffered solution (PBS) at a dose of about 1-5 mg/kg. The control consists of administering PBS only.

The procedure for testing the effects of the test compound would consist of intradermally injecting killed Mycobacterium tuberculosis in CFA followed by immediately administering the test compound and subsequent treatment every other day until day 24. At 14, 15, 18, 20, 22, and 24 days after injection of Mycobacterium CFA, an overall arthritis score may be obtained as described by J. Holoskitz above. An analysis of the data would reveal that the test compound would have a dramatic affect on the swelling of the joints as measured by a decrease of the arthritis score.

3.11 THERAPEUTIC METHODS

The compositions (including polypeptide fragments, analogs, variants and antibodies or other binding partners or modulators including antisense polynucleotides) of the invention have numerous applications in a variety of therapeutic methods. Examples of therapeutic applications include, but are not limited to, those exemplified herein.

3.11.1 EXAMPLE

One embodiment of the invention is the administration of an effective amount of the polypeptides or other composition of the invention to individuals affected by a disease or disorder that can be modulated by regulating the peptides of the invention. While the mode of administration is not particularly important, parenteral administration is preferred. An exemplary mode of administration is to deliver an intravenous bolus. The dosage of the polypeptides or other composition of the invention will normally be determined by the prescribing physician. It is to be expected that the dosage will vary according to the age, weight, condition and response of the individual patient. Typically, the amount of polypeptide administered per dose will be in the range of about 0.01μg/kg to 100 mg/kg of body weight, with the preferred dose being about 0.1μg/kg to 10 mg/kg of patient body weight. For parenteral administration, polypeptides of the invention will be formulated in an injectable form combined with a pharmaceutically acceptable parenteral vehicle. Such vehicles are well known in the art and examples include water, saline, Ringer's solution, dextrose solution, and solutions consisting of small amounts of the human serum albumin. The vehicle may contain minor amounts of additives that maintain the isotonicity and stability of the polypeptide or other active ingredient. The preparation of such solutions is within the skill of the art.

3.12 PHARMACEUTICAL FORMULATIONS AND ROUTES OF ADMINISTRATION

A protein or other composition of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources and including antibodies and other binding partners of the polypeptides of the invention) may be administered to a patient in need, by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s) at doses to treat or ameliorate a variety of disorders. Such a composition may optionally contain (in addition to protein or other active ingredient and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2,

G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the disease or disorder in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet-derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), insulin-like growth factor (IGF), as well as cytokines described herein.

The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or other active ingredient or complement its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein or other active ingredient of the invention, or to minimize side effects. Conversely, protein or other active ingredient of the present invention may be included in formulations of the particular clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent (such as IL-1Ra, IL-1 Hy1, IL-1 Hy2, anti-TNF, corticosteroids, immunosuppressive agents). A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

As an alternative to being included in a pharmaceutical composition of the invention including a first protein, a second protein or a therapeutic agent may be concurrently administered with the first protein (e.g., at the same time, or at differing times provided that therapeutic concentrations of the combination of agents is achieved at the treatment site). Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms, *e.g.*, treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, a therapeutically effective dose refers to that ingredient alone. When applied to a combination, a therapeutically effective dose refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein or other active ingredient of the present invention is administered to a mammal having a condition to be treated. Protein or other active ingredient of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co- administered with one or more cytokines, lymphokines or other hematopoietic factors, protein or other active ingredient of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein or other active ingredient of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

3.12.1 ROUTES OF ADMINISTRATION

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections. Administration of protein or other active ingredient of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into a arthritic joints or in fibrotic tissue, often in a depot or sustained release formulation. In order to prevent the scarring process frequently occurring as complication of glaucoma surgery, the compounds may be administered topically, for example, as eye drops. Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with a specific antibody, targeting, for example, arthritic or fibrotic tissue. The liposomes will be targeted to and taken up selectively by the afflicted tissue.

The polypeptides of the invention are administered by any route that delivers an effective dosage to the desired site of action. The determination of a suitable route of administration and an effective dosage for a particular indication is within the level of skill in

the art. Preferably for wound treatment, one administers the therapeutic compound directly to the site. Suitable dosage ranges for the polypeptides of the invention can be extrapolated from these dosages or from similar studies in appropriate animal models. Dosages can then be adjusted as necessary by the clinician to provide maximal therapeutic benefit.

3.12.2 COMPOSITIONS/FORMULATIONS

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in a conventional manner using one or more physiologically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. These pharmaceutical compositions may be manufactured in a manner that is itself known, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes. Proper formulation is dependent upon the route of administration chosen. When a therapeutically effective amount of protein or other active ingredient of the present invention is administered orally, protein or other active ingredient of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein or other active ingredient of the present invention, and preferably from about 25 to 90% protein or other active ingredient of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein or other active ingredient of the present invention, and preferably from about 1 to 50% protein or other active ingredient of the present invention.

When a therapeutically effective amount of protein or other active ingredient of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein or other active ingredient of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein or other active ingredient solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition

for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein or other active ingredient of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Pharmaceutical preparations for oral use can be obtained from a solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene

glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration. For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, *e.g.*, dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, *e.g.*, gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch. The compounds may be formulated for parenteral administration by injection, *e.g.*, by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, *e.g.*, in ampules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, *e.g.*, sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, *e.g.*, containing conventional suppository bases such as cocoa butter or other glycerides. In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable

polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a co-solvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The co-solvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD co-solvent system (VPD:5W) consists of VPD diluted 1:1 with a 5% dextrose in water solution. This co-solvent system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may replace polyethylene glycol, e.g. polyvinyl pyrrolidone; and other sugars or polysaccharides may substitute for dextrose. Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity. Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various types of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein or other active ingredient stabilization may be employed.

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols. Many of the active ingredients of the invention may be provided as salts with pharmaceutically compatible counter ions. Such pharmaceutically acceptable base addition salts are those salts which retain the biological effectiveness and properties of the free acids and which are obtained by reaction with inorganic or organic bases such as sodium hydroxide, magnesium hydroxide, ammonia,

trialkylamine, dialkylamine, monoalkylamine, dibasic amino acids, sodium acetate, potassium benzoate, triethanol amine and the like.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) or other active ingredient(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithins, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent Nos. 4,235,871; 4,501,728; 4,837,028; and 4,737,323, all of which are incorporated herein by reference.

The amount of protein or other active ingredient of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein or other active ingredient of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein or other active ingredient of the present invention and observe the patient's response. Larger doses of protein or other active ingredient of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 μg to about 100 mg

(preferably about 0.1 μg to about 10 mg, more preferably about 0.1 μg to about 1 mg) of protein or other active ingredient of the present invention per kg body weight. For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein or other active ingredient of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing or other active ingredient-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalcium phosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalcium phosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability. Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize

a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt %, preferably 1-10 wt % based on total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells. In further compositions, proteins or other active ingredients of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins or other active ingredients of the present invention. The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, *e.g.*, amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (*e.g.*, bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either in vivo or ex vivo into cells for expression in a

mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes.

3.12.3 EFFECTIVE DOSAGE

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount effective to prevent development of or to alleviate the existing symptoms of the subject being treated. Determination of the effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from appropriate in vitro assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that can be used to more accurately determine useful doses in humans. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC_{50} as determined in cell culture (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of the protein's biological activity). Such information can be used to more accurately determine useful doses in humans.

A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between LD_{50} and ED_{50}. Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of

administration and dosage can be chosen by the individual physician in view of the patient's condition. See, *e.g.*, Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1. Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety which are sufficient to maintain the desired effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from *in vitro* data. Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen which maintains plasma levels above the MEC for 10-90% of the time, preferably between 30-90% and most preferably between 50-90%. In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

An exemplary dosage regimen for polypeptides or other compositions of the invention will be in the range of about 0.01 μg/kg to 100 mg/kg of body weight daily, with the preferred dose being about 0.1 μg/kg to 25 mg/kg of patient body weight daily, varying in adults and children. Dosing may be once daily, or equivalent doses may be delivered at longer or shorter intervals.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's age and weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

3.12.4 PACKAGING

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may, for example, comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. Compositions comprising a compound of the invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition.

3.13 ANTIBODIES

Also included in the invention are antibodies to proteins, or fragments of proteins of the invention. The term "antibody" as used herein refers to immunoglobulin molecules and

immunologically active portions of immunoglobulin (Ig) molecules, i.e., molecules that contain an antigen-binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab}, $F_{ab'}$ and $F_{(ab')2}$ fragments, and an F_{ab} expression library. In general, an antibody molecule obtained from humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ from one another by the nature of the heavy chain present in the molecule. Certain classes have subclasses as well, such as IgG_1, IgG_2, and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all such classes, subclasses and types of human antibody species.

An isolated related protein of the invention may be intended to serve as an antigen, or a portion or fragment thereof, and additionally can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for polyclonal and monoclonal antibody preparation. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments of the antigen for use as immunogens. An antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein, such as an amino acid sequence shown in SEQ ID NO: 949-1896, and encompasses an epitope thereof such that an antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope. Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 15 amino acid residues, or at least 20 amino acid residues, or at least 30 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its surface; commonly these are hydrophilic regions.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of alpha-2-macroglobulin-like protein that is located on the surface of the protein, *e.g.*, a hydrophilic region. A hydrophobicity analysis of the human related protein sequence will indicate which regions of a related protein are particularly hydrophilic and, therefore, are likely to encode surface residues useful for targeting antibody production. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, *e.g.*, Hopp and Woods, 1981, *Proc. Nat. Acad. Sci. USA* 78: 3824-3828; Kyte and Doolittle 1982, *J. Mol. Biol.* 157: 105-142, each of which is incorporated herein by reference in its entirety. Antibodies that are specific for one or more

domains within an antigenic protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

A protein of the invention, or a derivative, fragment, analog, homolog or ortholog thereof, may be utilized as an immunogen in the generation of antibodies that immunospecifically bind these protein components.

The term "specific for" indicates that the variable regions of the antibodies of the invention recognize and bind polypeptides of the invention exclusively (*i.e.*, able to distinguish the polypeptide of the invention from other similar polypeptides despite sequence identity, homology, or similarity found in the family of polypeptides), but may also interact with other proteins (for example, *S. aureus* protein A or other antibodies in ELISA techniques) through interactions with sequences outside the variable region of the antibodies, and in particular, in the constant region of the molecule. Screening assays to determine binding specificity of an antibody of the invention are well known and routinely practiced in the art. For a comprehensive discussion of such assays, see Harlow et al. (Eds), Antibodies A Laboratory Manual; Cold Spring Harbor Laboratory; Cold Spring Harbor, NY (1988), Chapter 6. Antibodies that recognize and bind fragments of the polypeptides of the invention are also contemplated, provided that the antibodies are first and foremost specific for, as defined above, full-length polypeptides of the invention. As with antibodies that are specific for full length polypeptides of the invention, antibodies of the invention that recognize fragments are those which can distinguish polypeptides from the same family of polypeptides despite inherent sequence identity, homology, or similarity found in the family of proteins.

Antibodies of the invention are useful for, for example, therapeutic purposes (by modulating activity of a polypeptide of the invention), diagnostic purposes to detect or quantitate a polypeptide of the invention, as well as purification of a polypeptide of the invention. Kits comprising an antibody of the invention for any of the purposes described herein are also comprehended. In general, a kit of the invention also includes a control antigen for which the antibody is immunospecific. The invention further provides a hybridoma that produces an antibody according to the invention. Antibodies of the invention are useful for detection and/or purification of the polypeptides of the invention.

Monoclonal antibodies binding to the protein of the invention may be useful diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells,

neutralizing monoclonal antibodies against the protein may be useful in detecting and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein.

The labeled antibodies of the present invention can be used for *in vitro*, *in vivo*, and *in situ* assays to identify cells or tissues in which a fragment of the polypeptide of interest is expressed. The antibodies may also be used directly in therapies or other diagnostics. The present invention further provides the above-described antibodies immobilized on a solid support. Examples of such solid supports include plastics such as polycarbonate, complex carbohydrates such as agarose and Sepharose®, acrylic resins and such as polyacrylamide and latex beads. Techniques for coupling antibodies to such solid supports are well known in the art (Weir, D.M. et al., "Handbook of Experimental Immunology" 4th Ed., Blackwell Scientific Publications, Oxford, England, Chapter 10 (1986); Jacoby, W.D. et al., Meth. Enzym. 34 Academic Press, N.Y. (1974)). The immobilized antibodies of the present invention can be used for *in vitro*, *in vivo*, and *in situ* assays as well as for immuno-affinity purification of the proteins of the present invention.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, fragments, analogs homologs or orthologs thereof (see, for example, Antibodies: A Laboratory Manual, Harlow E, and Lane D, 1988, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

3.13.1 POLYCLONAL ANTIBODIES

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the immunogenic protein, or a recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g.,

aluminum hydroxide), surface-active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), adjuvants usable in humans such as Bacille Calmette-Guerin and Corynebacterium parvum, or similar immunostimulatory agents. Additional examples of adjuvants that can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate).

The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (The Scientist, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

3.13.2 MONOCLONAL ANTIBODIES

The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one molecular species of antibody molecule consisting of a unique light chain gene product and a unique heavy chain gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal antibody are identical in all the molecules of the population. MAbs thus contain an antigen-binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

Monoclonal antibodies can be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be immunized in vitro.

The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-

103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980). Preferably, antibodies having a high degree of specificity and a high binding affinity for the target antigen are isolated.

After the desired hybridoma cells are identified, the clones can be subcloned by limiting dilution procedures and grown by standard methods. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal.

The monoclonal antibodies secreted by the subclones can be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA can be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, Nature 368, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

3.13.3 HUMANIZED ANTIBODIES

The antibodies directed against the protein antigens of the invention can further comprise humanized antibodies or human antibodies. These antibodies are suitable for administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', $F(ab')_2$ or other antigen-binding subsequences of antibodies) that are principally comprised of the sequence of a human immunoglobulin, and contain minimal sequence derived from a non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeyen et al., Science, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. (See also U.S. Patent No. 5,225,539). In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues that are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise

substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., 1986; Riechmann et al., 1988; and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)).

3.13.4 HUMAN ANTIBODIES

Fully human antibodies relate to antibody molecules in which essentially the entire sequences of both the light chain and the heavy chain, including the CDRs, arise from human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein. Human monoclonal antibodies can be prepared by the trioma technique; the human B-cell hybridoma technique (see Kozbor, et al., 1983 Immunol Today 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, et al., 1983. Proc Natl Acad Sci USA 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96).

In addition, human antibodies can also be produced using additional techniques, including phage display libraries (Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks et al. (Bio/Technology 10, 779-783 (1992)); Lonberg et al. (Nature 368 856-859 (1994)); Morrison (Nature 368, 812-13 (1994)); Fishwild et al, (Nature Biotechnology 14, 845-51 (1996)); Neuberger (Nature Biotechnology 14, 826 (1996)); and Lonberg and Huszar (Intern. Rev. Immunol. 13 65-93 (1995)).

Human antibodies may additionally be produced using transgenic nonhuman animals that are modified so as to produce fully human antibodies rather than the animal's endogenous antibodies in response to challenge by an antigen. (See PCT publication WO94/02602). The endogenous genes encoding the heavy and light immunoglobulin chains in the nonhuman host have been incapacitated, and active loci encoding human heavy and light chain immunoglobulins are inserted into the host's genome. The human genes are incorporated, for example, using yeast artificial chromosomes containing the requisite human DNA segments. An animal which provides all the desired modifications is then obtained as progeny by crossbreeding intermediate transgenic animals containing fewer than the full complement of the modifications. The preferred embodiment of such a nonhuman animal is a mouse, and is termed the Xenomouse™ as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells that secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as hybridomas producing monoclonal antibodies. Additionally, the genes encoding the immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S. Patent No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into another mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

3.13.5 FAB FRAGMENTS AND SINGLE CHAIN ANTIBODIES

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see e.g., U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression libraries (see e.g., Huse, et al., 1989 Science 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an $F_{(ab')2}$ fragment produced by pepsin digestion of an antibody molecule; (ii) an F_{ab} fragment generated by reducing the disulfide bridges of an $F_{(ab')2}$ fragment; (iii) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F_v fragments.

3.13.6 BISPECIFIC ANTIBODIES

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, Nature, 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, 1991 *EMBO J.*, 10:3655-3659.

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion

preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers that are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. $F(ab')_2$ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., Science 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate $F(ab')_2$ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally, Fab' fragments can be directly recovered from E. coli and chemically coupled to form bispecific antibodies. Shalaby et al., J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody $F(ab')_2$ molecule. Each Fab' fragment was separately secreted from E. coli and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was

able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., J. Immunol. 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., J. Immunol. 147:60 (1991).

Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (FcγR), such as FcγRI (CD64), FcγRII (CD32) and FcγRIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular antigen. Bispecific antibodies can also be used to direct cytotoxic agents to cells which express a particular antigen. These antibodies possess an antigen-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the protein antigen described herein and further binds tissue factor (TF).

3.13.7 HETEROCONJUGATE ANTIBODIES

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins can be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

3.13.8 EFFECTOR FUNCTION ENGINEERING

It can be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated can have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff et al. Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and can thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design, 3: 219-230 (1989).

3.13.9 IMMUNOCONJUGATES

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from Pseudomonas aeruginosa), ricin A chain, abrin A chain, modeccin A chain,

alpha-sarcin, Aleurites fordii proteins, dianthin proteins, Phytolaca americana proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ^{212}Bi, ^{131}I, ^{131}In, ^{90}Y, and ^{186}Re.

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody can be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is in turn conjugated to a cytotoxic agent.

3.14 COMPUTER READABLE SEQUENCES

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium which can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention. As used herein, "recorded" refers to a process for storing information on computer

readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (*e.g.* text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing any of the nucleotide sequences SEQ ID NOs: 1 - 948 or a representative fragment thereof; or a nucleotide sequence at least 95% identical to any of the nucleotide sequences of SEQ ID NOs: 1 - 948 in computer readable form, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul et al., J. Mol. Biol. 215:403-410 (1990)) and BLAZE (Brutlag et al., Comp. Chem. 17:203-207 (1993)) search algorithms on a Sybase system is used to identify open reading frames (ORFs) within a nucleic acid sequence. Such ORFs may be protein encoding fragments and may be useful in producing commercially important proteins such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based systems are suitable for use in the present invention. As stated above, the computer-based systems of the present invention comprise a data storage means

having stored therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory which can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of a known sequence which match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but is not limited to, Smith-Waterman, MacPattern (EMBL), BLASTN and BLASTA (NPOLYPEPTIDEIA). A skilled artisan can readily recognize that any one of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems. As used herein, a "target sequence" can be any nucleic acid or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 300 amino acids, more preferably from about 30 to 100 nucleotide residues. However, it is well recognized that searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

3.15 TRIPLE HELIX FORMATION

In addition, the fragments of the present invention, as broadly described, can be used to control gene expression through triple helix formation or antisense DNA or RNA, both of which methods are based on the binding of a polynucleotide sequence to DNA or RNA. Polynucleotides suitable for use in these methods are preferably 20 to 40 bases in length and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 15241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Olmno, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide.

3.16 DIAGNOSTIC ASSAYS AND KITS

The present invention further provides methods to identify the presence or expression of one of the ORFs of the present invention, or homolog thereof, in a test sample, using a nucleic acid probe or antibodies of the present invention, optionally conjugated or otherwise associated with a suitable label.

In general, methods for detecting a polynucleotide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polynucleotide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polynucleotide of the invention is detected in the sample. Such methods can also comprise contacting a sample under stringent hybridization conditions with nucleic acid primers that anneal to a polynucleotide of the invention under such conditions, and amplifying annealed polynucleotides, so that if a polynucleotide is amplified, a polynucleotide of the invention is detected in the sample.

In general, methods for detecting a polypeptide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polypeptide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polypeptide of the invention is detected in the sample.

In detail, such methods comprise incubating a test sample with one or more of the antibodies or one or more of the nucleic acid probes of the present invention and assaying for binding of the nucleic acid probes or antibodies to components within the test sample.

Conditions for incubating a nucleic acid probe or antibody with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid probe or antibody used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization, amplification or immunological assay formats can readily be adapted to employ the nucleic acid probes or antibodies of the present invention. Examples of such assays can be found in Chard, T., An Introduction to Radioimmunoassay and Related Techniques, Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G.R. et al., Techniques in Immunocytochemistry, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., Practice and Theory of immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1985). The test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention. Specifically, the invention provides a compartment kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the probes or antibodies of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound probe or antibody.

In detail, a compartment kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the antibodies used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound

antibody or probe. Types of detection reagents include labeled nucleic acid probes, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. One skilled in the art will readily recognize that the disclosed probes and antibodies of the present invention can be readily incorporated into one of the established kit formats which are well known in the art.

3.17 MEDICAL IMAGING

The novel polypeptides and binding partners of the invention are useful in medical imaging of sites expressing the molecules of the invention (e.g., where the polypeptide of the invention is involved in the immune response, for imaging sites of inflammation or infection). See, e.g., Kunkel et al., U.S. Pat. NO. 5,413,778. Such methods involve chemical attachment of a labeling or imaging agent, administration of the labeled polypeptide to a subject in a pharmaceutically acceptable carrier, and imaging the labeled polypeptide *in vivo* at the target site.

3.18 SCREENING ASSAYS

Using the isolated proteins and polynucleotides of the invention, the present invention further provides methods of obtaining and identifying agents which bind to a polypeptide encoded by an ORF corresponding to any of the nucleotide sequences set forth in SEQ ID NOs: 1 - 948, or bind to a specific domain of the polypeptide encoded by the nucleic acid. In detail, said method comprises the steps of:

(a) contacting an agent with an isolated protein encoded by an ORF of the present invention, or nucleic acid of the invention; and

(b) determining whether the agent binds to said protein or said nucleic acid.

In general, therefore, such methods for identifying compounds that bind to a polynucleotide of the invention can comprise contacting a compound with a polynucleotide of the invention for a time sufficient to form a polynucleotide/compound complex, and detecting the complex, so that if a polynucleotide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Likewise, in general, therefore, such methods for identifying compounds that bind to a polypeptide of the invention can comprise contacting a compound with a polypeptide of the invention for a time sufficient to form a polypeptide/compound complex, and detecting the

complex, so that if a polypeptide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Methods for identifying compounds that bind to a polypeptide of the invention can also comprise contacting a compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a receptor gene sequence in the cell, and detecting the complex by detecting reporter gene sequence expression, so that if a polypeptide/compound complex is detected, a compound that binds a polypeptide of the invention is identified.

Compounds identified via such methods can include compounds which modulate the activity of a polypeptide of the invention (that is, increase or decrease its activity, relative to activity observed in the absence of the compound). Alternatively, compounds identified via such methods can include compounds which modulate the expression of a polynucleotide of the invention (that is, increase or decrease expression relative to expression levels observed in the absence of the compound). Compounds, such as compounds identified via the methods of the invention, can be tested using standard assays well known to those of skill in the art for their ability to modulate activity/expression.

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be selected and screened at random or rationally selected or designed using protein modeling techniques.

For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein encoded by the ORF of the present invention. Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular protein. For example, one skilled in the art can readily adapt currently available procedures to generate peptides, pharmaceutical agents and the like, capable of binding to a specific peptide sequence, in order to generate rationally designed antipeptide peptides, for example see Hurby et al., Application of Synthetic Peptides: Antisense Peptides," In Synthetic Peptides, A User's Guide, W.H. Freeman, NY (1992), pp. 289-307, and Kaspczak et al., Biochemistry 28:9230-8 (1989), or pharmaceutical agents, or the like.

In addition to the foregoing, one class of agents of the present invention, as broadly described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or

rationally designed/selected. Targeting the ORF or EMF allows a skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or multiple ORFs which rely on the same EMF for expression control. One class of DNA binding agents are agents which contain base residues which hybridize or form a triple helix formation by binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or can be a variety of sulfhydryl or polymeric derivatives which have base attachment capacity.

Agents suitable for use in these methods preferably contain 20 to 40 bases and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide and other DNA binding agents.

Agents which bind to a protein encoded by one of the ORFs of the present invention can be used as a diagnostic agent. Agents which bind to a protein encoded by one of the ORFs of the present invention can be formulated using known techniques to generate a pharmaceutical composition.

3.19 USE OF NUCLEIC ACIDS AS PROBES

Another aspect of the subject invention is to provide for polypeptide-specific nucleic acid hybridization probes capable of hybridizing with naturally occurring nucleotide sequences. The hybridization probes of the subject invention may be derived from any of the nucleotide sequences SEQ ID NOs: 1 - 948. Because the corresponding gene is only expressed in a limited number of tissues, a hybridization probe derived from of any of the nucleotide sequences SEQ ID NOs: 1 - 948 can be used as an indicator of the presence of RNA of cell type of such a tissue in a sample.

Any suitable hybridization technique can be employed, such as, for example, in situ hybridization. PCR as described in US Patents Nos. 4,683,195 and 4,965,188 provides additional uses for oligonucleotides based upon the nucleotide sequences. Such probes used

in PCR may be of recombinant origin, may be chemically synthesized, or a mixture of both. The probe will comprise a discrete nucleotide sequence for the detection of identical sequences or a degenerate pool of possible sequences for identification of closely related genomic sequences.

Other means for producing specific hybridization probes for nucleic acids include the cloning of nucleic acid sequences into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled nucleotides. The nucleotide sequences may be used to construct hybridization probes for mapping their respective genomic sequences. The nucleotide sequence provided herein may be mapped to a chromosome or specific regions of a chromosome using well known genetic and/or chromosomal mapping techniques. These techniques include in situ hybridization, linkage analysis against known chromosomal markers, hybridization screening with libraries or flow-sorted chromosomal preparations specific to known chromosomes, and the like. The technique of fluorescent in situ hybridization of chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY.

Fluorescent *in situ* hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of a nucleic acid on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.

3.20 PREPARATION OF SUPPORT BOUND OLIGONUCLEOTIDES

Oligonucleotides, i.e., small nucleic acid segments, may be readily prepared by, for example, directly synthesizing the oligonucleotide by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer.

Support bound oligonucleotides may be prepared by any of the methods known to those of skill in the art using any suitable support such as glass, polystyrene or Teflon. One strategy is to precisely spot oligonucleotides synthesized by standard synthesizers. Immobilization can be

achieved using passive adsorption (Inouye & Hondo, (1990) J. Clin. Microbiol. 28(6) 1469-72); using UV light (Nagata *et al.*, 1985; Dahlen *et al.*, 1987; Morrissey & Collins, (1989) Mol. Cell Probes 3(2) 189-207) or by covalent binding of base modified DNA (Keller *et al.*, 1988; 1989); all references being specifically incorporated herein.

Another strategy that may be employed is the use of the strong biotin-streptavidin interaction as a linker. For example, Broude *et al.* (1994) Proc. Natl. Acad. Sci. USA 91(8) 3072-6, describe the use of biotinylated probes, although these are duplex probes, that are immobilized on streptavidin-coated magnetic beads. Streptavidin-coated beads may be purchased from Dynal, Oslo. Of course, this same linking chemistry is applicable to coating any surface with streptavidin. Biotinylated probes may be purchased from various sources, such as, e.g., Operon Technologies (Alameda, CA).

Nunc Laboratories (Naperville, IL) is also selling suitable material that could be used. Nunc Laboratories have developed a method by which DNA can be covalently bound to the microwell surface termed Covalink NH. CovaLink NH is a polystyrene surface grafted with secondary amino groups (>NH) that serve as bridge-heads for further covalent coupling. CovaLink Modules may be purchased from Nunc Laboratories. DNA molecules may be bound to CovaLink exclusively at the 5'-end by a phosphoramidate bond, allowing immobilization of more than 1 pmol of DNA (Rasmussen *et al.*, (1991) Anal. Biochem. 198(1) 138-42).

The use of CovaLink NH strips for covalent binding of DNA molecules at the 5'-end has been described (Rasmussen et al., (1991). In this technology, a phosphoramidate bond is employed (Chu et al., (1983) Nucleic Acids Res. 11(8) 6513-29). This is beneficial as immobilization using only a single covalent bond is preferred. The phosphoramidate bond joins the DNA to the CovaLink NH secondary amino groups that are positioned at the end of spacer arms covalently grafted onto the polystyrene surface through a 2 nm long spacer arm. To link an oligonucleotide to CovaLink NH via an phosphoramidate bond, the oligonucleotide terminus must have a 5'-end phosphate group. It is, perhaps, even possible for biotin to be covalently bound to CovaLink and then streptavidin used to bind the probes.

More specifically, the linkage method includes dissolving DNA in water (7.5 ng/ul) and denaturing for 10 min. at 95°C and cooling on ice for 10 min. Ice-cold 0.1 M 1-methylimidazole, pH 7.0 (1-MeIm$_7$), is then added to a final concentration of 10 mM 1-MeIm$_7$. A ss DNA solution is then dispensed into CovaLink NH strips (75 ul/well) standing on ice.

Carbodiimide 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC), dissolved in 10 mM 1-MeIm$_7$, is made fresh and 25 ul added per well. The strips are incubated for 5 hours

at 50°C. After incubation the strips are washed using, e.g., Nunc-Immuno Wash; first the wells are washed 3 times, then they are soaked with washing solution for 5 min., and finally they are washed 3 times (where in the washing solution is 0.4 N NaOH, 0.25% SDS heated to 50°C).

It is contemplated that a further suitable method for use with the present invention is that described in PCT Patent Application WO 90/03382 (Southern & Maskos), incorporated herein by reference. This method of preparing an oligonucleotide bound to a support involves attaching a nucleoside 3'-reagent through the phosphate group by a covalent phosphodiester link to aliphatic hydroxyl groups carried by the support. The oligonucleotide is then synthesized on the supported nucleoside and protecting groups removed from the synthetic oligonucleotide chain under standard conditions that do not cleave the oligonucleotide from the support. Suitable reagents include nucleoside phosphoramidite and nucleoside hydrogen phosphorate.

An on-chip strategy for the preparation of DNA probe for the preparation of DNA probe arrays may be employed. For example, addressable laser-activated photodeprotection may be employed in the chemical synthesis of oligonucleotides directly on a glass surface, as described by Fodor *et al.* (1991) Science 251(4995) 767-73, incorporated herein by reference. Probes may also be immobilized on nylon supports as described by Van Ness *et al.* (1991) Nucleic Acids Res. 19(12) 3345-50; or linked to Teflon using the method of Duncan & Cavalier (1988) Anal. Biochem. 169(1) 104-8; all references being specifically incorporated herein.

To link an oligonucleotide to a nylon support, as described by Van Ness *et al.* (1991), requires activation of the nylon surface via alkylation and selective activation of the 5'-amine of oligonucleotides with cyanuric chloride.

One particular way to prepare support bound oligonucleotides is to utilize the light-generated synthesis described by Pease *et al.*, (1994) PNAS USA 91(11) 5022-6, incorporated herein by reference). These authors used current photolithographic techniques to generate arrays of immobilized oligonucleotide probes (DNA chips). These methods, in which light is used to direct the synthesis of oligonucleotide probes in high-density, miniaturized arrays, utilize photolabile 5'-protected *N*-acyl-deoxynucleoside phosphoramidites, surface linker chemistry and versatile combinatorial synthesis strategies. A matrix of 256 spatially defined oligonucleotide probes may be generated in this manner.

3.21 PREPARATION OF NUCLEIC ACID FRAGMENTS

The nucleic acids may be obtained from any appropriate source, such as cDNAs, genomic DNA, chromosomal DNA, microdissected chromosome bands, cosmid or YAC inserts, and RNA, including mRNA without any amplification steps. For example, Sambrook *et al.*

(1989) describes three protocols for the isolation of high molecular weight DNA from mammalian cells (p. 9.14-9.23).

DNA fragments may be prepared as clones in M13, plasmid or lambda vectors and/or prepared directly from genomic DNA or cDNA by PCR or other amplification methods. Samples may be prepared or dispensed in multiwell plates. About 100-1000 ng of DNA samples may be prepared in 2-500 ml of final volume.

The nucleic acids would then be fragmented by any of the methods known to those of skill in the art including, for example, using restriction enzymes as described at 9.24-9.28 of Sambrook *et al.* (1989), shearing by ultrasound and NaOH treatment.

Low pressure shearing is also appropriate, as described by Schriefer *et al.* (1990) Nucleic Acids Res. 18(24) 7455-6, incorporated herein by reference). In this method, DNA samples are passed through a small French pressure cell at a variety of low to intermediate pressures. A lever device allows controlled application of low to intermediate pressures to the cell. The results of these studies indicate that low-pressure shearing is a useful alternative to sonic and enzymatic DNA fragmentation methods.

One particularly suitable way for fragmenting DNA is contemplated to be that using the two base recognition endonuclease, *Cvi*JI, described by Fitzgerald *et al.* (1992) Nucleic Acids Res. 20(14) 3753-62. These authors described an approach for the rapid fragmentation and fractionation of DNA into particular sizes that they contemplated to be suitable for shotgun cloning and sequencing.

The restriction endonuclease *Cvi*JI normally cleaves the recognition sequence PuGCPy between the G and C to leave blunt ends. Atypical reaction conditions, which alter the specificity of this enzyme (*Cvi*JI**), yield a quasi-random distribution of DNA fragments form the small molecule pUC19 (2688 base pairs). Fitzgerald *et al.* (1992) quantitatively evaluated the randomness of this fragmentation strategy, using a *Cvi*JI** digest of pUC19 that was size fractionated by a rapid gel filtration method and directly ligated, without end repair, to a lac Z minus M13 cloning vector. Sequence analysis of 76 clones showed that *Cvi*JI** restricts pyGCPy and PuGCPu, in addition to PuGCPy sites, and that new sequence data is accumulated at a rate consistent with random fragmentation.

As reported in the literature, advantages of this approach compared to sonication and agarose gel fractionation include: smaller amounts of DNA are required (0.2-0.5 ug instead of 2-5 ug); and fewer steps are involved (no preligation, end repair, chemical extraction, or agarose gel electrophoresis and elution are needed

Irrespective of the manner in which the nucleic acid fragments are obtained or prepared, it is important to denature the DNA to give single stranded pieces available for hybridization. This is achieved by incubating the DNA solution for 2-5 minutes at 80-90°C. The solution is then cooled quickly to 2°C to prevent renaturation of the DNA fragments before they are contacted with the chip. Phosphate groups must also be removed from genomic DNA by methods known in the art.

3.22 PREPARATION OF DNA ARRAYS

Arrays may be prepared by spotting DNA samples on a support such as a nylon membrane. Spotting may be performed by using arrays of metal pins (the positions of which correspond to an array of wells in a microtiter plate) to repeated by transfer of about 20 nl of a DNA solution to a nylon membrane. By offset printing, a density of dots higher than the density of the wells is achieved. One to 25 dots may be accommodated in 1 mm^2, depending on the type of label used. By avoiding spotting in some preselected number of rows and columns, separate subsets (subarrays) may be formed. Samples in one subarray may be the same genomic segment of DNA (or the same gene) from different individuals, or may be different, overlapped genomic clones. Each of the subarrays may represent replica spotting of the same samples. In one example, a selected gene segment may be amplified from 64 patients. For each patient, the amplified gene segment may be in one 96-well plate (all 96 wells containing the same sample). A plate for each of the 64 patients is prepared. By using a 96-pin device, all samples may be spotted on one 8 x 12 cm membrane. Subarrays may contain 64 samples, one from each patient. Where the 96 subarrays are identical, the dot span may be 1 mm^2 and there may be a 1 mm space between subarrays.

Another approach is to use membranes or plates (available from NUNC, Naperville, Illinois) which may be partitioned by physical spacers e.g. a plastic grid molded over the membrane, the grid being similar to the sort of membrane applied to the bottom of multiwell plates, or hydrophobic strips. A fixed physical spacer is not preferred for imaging by exposure to flat phosphor-storage screens or x-ray films.

The present invention is illustrated in the following examples. Upon consideration of the present disclosure, one of skill in the art will appreciate that many other embodiments and variations may be made in the scope of the present invention. Accordingly, it is intended that the broader aspects of the present invention not be limited to the disclosure of the following examples. The present invention is not to be limited in scope by the exemplified embodiments which are intended as illustrations of single aspects of the invention, and compositions and

methods which are functionally equivalent are within the scope of the invention. Indeed, numerous modifications and variations in the practice of the invention are expected to occur to those skilled in the art upon consideration of the present preferred embodiments. Consequently, the only limitations which should be placed upon the scope of the invention are those which appear in the appended claims.

All references cited within the body of the instant specification are hereby incorporated by reference in their entirety.

4.0 EXAMPLES

4.1 EXAMPLE 1

Novel Nucleic Acid Sequences Obtained From Various Libraries

A plurality of novel nucleic acids were obtained from cDNA libraries prepared from various human tissues and in some cases isolated from a genomic library derived from human chromosome using standard PCR, SBH sequence signature analysis and Sanger sequencing techniques. The inserts of the library were amplified with PCR using primers specific for the vector sequences which flank the inserts. Clones from cDNA libraries were spotted on nylon membrane filters and screened with oligonucleotide probes (e.g., 7-mers) to obtain signature sequences. The clones were clustered into groups of similar or identical sequences. Representative clones were selected for sequencing.

In some cases, the 5' sequence of the amplified inserts was then deduced using a typical Sanger sequencing protocol. PCR products were purified and subjected to fluorescent dye terminator cycle sequencing. Single pass gel sequencing was done using a 377 Applied Biosystems (ABI) sequencer to obtain the novel nucleic acid sequences. In some cases RACE (Random Amplification of cDNA Ends) was performed to further extend the sequence in the 5' direction.

4.2 EXAMPLE 2

Novel Nucleic Acids

The novel nucleic acids of the present invention of the invention were assembled from sequences that were obtained from a cDNA library by methods described in Example 1 above, and in some cases sequences obtained from one or more public databases. The nucleic acids were assembled using an EST sequence as a seed. Then a recursive algorithm was used to extend the seed EST into an extended assemblage, by pulling additional sequences from different databases (i.e., Hyseq's database containing EST sequences, dbEST version 119, gb pri

119, and UniGene version 119) that belong to this assemblage. The algorithm terminated when there was no additional sequences from the above databases that would extend the assemblage. Inclusion of component sequences into the assemblage was based on a BLASTN hit to the extending assemblage with BLAST score greater than 300 and percent identity greater than 95%.

Using PHRAP (Univ. of Washington) or CAP4 (Paracel), a full length gene cDNA sequence and its corresponding protein sequence were generated from the assemblage. Any frame shifts and incorrect stop codons were corrected by hand editing. During editing, the sequence was checked using FASTY and/or BLAST against Genbank (i.e., dbEST version 121, gb pri 121, UniGene version 121, Genpept release 121). Other computer programs which may have been used in the editing process were phredPhrap and Consed (University of Washington) and ed-ready, ed-ext and cg-zip-2 (Hyseq, Inc.). The full-length nucleotide and amino acid sequences, including splice variants resulting from these procedures are shown in the Sequence Listing as SEQ ID NOS: 1- 948.

Table 1 shows the various tissue sources of SEQ ID NO: 1-1896.

The nearest neighbor results for polypeptides encoded by SEQ ID NO: 1-948 (i.e. SEQ ID NO: 949-1896) were obtained by a BLASTP (version 2.0a1 19MP-WashU) search against Genpept, Geneseq and SwissProt databases using BLAST algorithm. The nearest neighbor result showed the closest homologue with functional annotation for SEQ ID NO: 1-948. The translated amino acid sequences for which the nucleic acid sequence encodes are shown in the Sequence Listing. The homologues with identifiable functions for SEQ ID NO: 1-948 are shown in Table 2 below.

Using eMatrix software package (Stanford University, Stanford, CA) (Wu et al., J. Comp. Biol., Vol. 6 pp. 219-235 (1999) herein incorporated by reference), polypeptides encoded by SEQ ID NO: 1-948 (i.e. SEQ ID NO: 949-1896) were examined to determine whether they had identifiable signature regions. Table 3 shows the signature region found in the indicated polypeptide sequences, the description of the signature, the eMatrix p-value(s) and the position(s) of the signature within the polypeptide sequence.

Using the Pfam software program (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1) pp. 320-322 (1998) herein incorporated by reference) polypeptides encoded by SEQ ID NO: 1-948 (i.e. SEQ ID NO: 949-1896) were examined for domains with homology to certain peptide domains. Table 4 shows the name of the domain found, the description, the product of all the e-value of similar domains found, the pFam score for the identified domain within

the sequence, number of similar domains found, and the position of the domain in the SEQ ID NO: being interrogated.

The GeneAtlas™ software package (Molecular Simulations Inc. (MSI), San Diego, CA) was used to predict the three-dimensional structure models for the polypeptides encoded by SEQ ID NO: 1-948 (i.e. SEQ ID NO: 949-1896). Models were generated by (1) PSI-BLAST which is a multiple alignment sequence profile-based searching developed by Altschul et al., (Nucl. Acids res. 25, 3389-3408 (1997)), (2) High Throughput Modeling (HTM) (Molecular Simulations Inc. (MSI) San Diego, CA) which is an automated sequence and structure searching procedure (http://www.msi.com/), and (3) SeqFold™ which is a fold recognition method described by Fischer and Eisenberg (J. Mol. Biol. 209, 779-791 (1998)). This analysis was carried out, in part, by comparing the polypeptides of the invention with the known NMR (nuclear magnetic resonance) and x-ray crystal three-dimensional structures as templates. Table 5 shows, "PDB ID", the Protein DataBase (PDB) identifier given to template structure; "Chain ID", identifier of the subcomponent of the PDB template structure; "Compound Information", information of the PDB template structure and/or its subcomponents; "PDB Function Annotation" gives function of the PDB template as annotated by the PDB files (http://www.rcsb.org/PDB/); start and end amino acid position of the protein sequence aligned; PSI-BLAST score, the verify score, the SeqFold score, and the Potential(s) of Mean Force (PMF). The verify score produced by GeneAtlas™ software (MSI), is based on Dr. Eisenberg's Profile-3D threading program developed in Dr. DavidEisenberg's laboratory (US patent no. 5,436,850 and Luthy, Bowie, and Eisenberg, Nature, 356:83-85 (1992)) and a publication by R. Sanchez and A. Sali, Proc. Natl. Acad. Sci. USA, 95:12502-13597. The verify score produced by GeneAtlas™ normalizes the verify score for proteins with different lengths so that a unified cutoff can be used to select good models as follows:

Verify score (normalized) = (raw score – ½ high score)/(1/2 high score)

The PMF score, produced by GeneAtlas™ software (MSI), is a composite scoring function that depends in part on the compactness of the model, sequence identity in the alignment used to build the model, pairwise and surface mean force potential (MFP). As given in Table 5, a verify score between 0 to 1.0, with 1 being the best, represents a good model. Similarly, a PMF score between 0 to 1.0, with 1 being the best, represents a good model. A SeqFold™ score of more than 50 is considered significant. A good model may

also be determined by one of skill in the art based on all the information in Table 5 taken in totality.

The nucleotide sequence within the sequences that codes for signal peptide sequences and their cleavage sites can be determined from using Neural Network SignalP V1.1 program(from Center for Biological Sequence Analysis, The Technical University of Denmark). The process for identifying prokaryotic and eukaryotic signal peptides and their cleavage sites are also disclosed by Henrik Nielson, Jacob Engelbrecht, Soren Brunak, and Gunnar von Heijne in the publication "Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites" Protein Engineering, Vol. 10, no. 1, pp. 1-6 (1997), incorporated herein by reference. A maximum S score and a mean S score, as described in the Nielson et al., as reference, were obtained for the polypeptide sequences. Table 6 shows the position of the last amino acid of the signal peptide in each of the polypeptides and the maximum score and mean score associated with that signal peptide.

Table 7 correlates each of SEQ ID NO: 1-948 to a specific chromosomal location.

Table 8 is a correlation table of the novel polynucleotide sequences SEQ ID NO: 1-948, novel polypeptide sequences SEQ ID NO: 949-1896, and their corresponding priority nucleotide sequences in the priority application USSN 09/799,451, herein incorporated by reference in its entirety.

TABLE 1

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
Null	mix tissues library	CTL016	52 137 189-192 316-325 529 591
Null	enriched libray	CTL021	65 84 169 189-192 311 316-325 406 676 727 782 850
Null	mix tissues library	CTL028	65
Null	PCR products cloning	PCR2V1	34 51 134 189 191-192 224 232 260 311 388 606 623 806
Null	mix tissues library	SUP002	51 96 103 163 216 272 294 311 316-325 328 378 383 388 446-448 450 453 474 481 500 516 610 774 780 885 904 922
adipocytes	Stratagene	ADP001	2 43 51 73 76 88 97 142 166 181 186 188 208 257 262-263 267-270 282 311 316-325 383 386 427-429 459 463 465 493 507 514 522 545 552 572 643 651 667 700 721 740 754 758 778 795 872 881 883 888 947
adrenal gland	Clontech	ADR002	3-6 10-11 13 16 20-21 24 27-28 33 38 48-49 51 53-54 58 66-67 75 88 97 99 124-125 130 140 157-158 179 188 197-198 200 212-214 216 218 224 229 231 237 257 267 279 281-282 288 302 311 326 362 376-377 381 383 396 398-403 429 443 453-454 456 459-460 474 489 515 526 531-

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
			532 540 545 550 559 564 568 577 581 586 589 599 605 610 613 631-632 643 648 651 667 670 672 681 684 699 703 706 708 717-718 734 736 751 779 785-786 795 813 817 837 871 876 887-888 897 904 907 916 921 924-926 948
adult brain	BioChain	ABR012	140 208 311 748 810
adult brain	BioChain	ABR013	51 245 311 316-325 436 717 810 936
adult brain	Clontech	ABR001	12 51 87 142 169 178 180 245 263 286 288 290 295 304 308 311-313 375 379-380 403 425 428 431 458 486 499 503 512 557-558 567-568 606 610 641 651 695 704 730 741 754 766 810 822 827 841 850 864 871 884 897 917 920 925-927 934 946
adult brain	Clontech	ABR006	2 14-15 22-23 29 32-33 49 66-68 83 99 111-112 115 129 131 142 147 153 157 163 169 189-192 200 205 207 212-214 218 221 229 234 256-257 263 272 276 279 282 292-299 301 311 315 340-343 349 376-377 383-386 388 403 405 407 410 425 438 453-454 460 463 469 474 489 495 499-500 511 522 531-532 539 541 545-546 551 556 563 565 571 579-583 591 594 606 626 628 631-632 643 647 651 678 684 691-692 700 717 721 726 730 732 741 744 754 757 769 772 774 782 788 793 810 820 827-828 853 867 869 875 879 897 913 921-922 925-926 933-934 939-941 947
adult brain	Clontech	ABR008	1-2 9-10 13 16-18 23 27-28 30-32 37 39 42-43 46 49-51 66 70 76 80 83 86-87 95-97 109 111-112 116-117 124 130-131 133-134 136-137 141-142 146-147 152-157 160 162 169 171 179 184 189-192 195 200-201 206 211-212 216-218 239 247-248 250 252 254-258 261-263 271-272 276 278 282 288 293-295 297 300 302 307 309 311 314-326 328 333 337-341 343 347 349 351-354 358 360-361 367 374 376-378 381 384 388-390 393 395-396 400-403 405 407 409 411 414 418-420 422 427-429 433 438 440-441 445-447 450 453-455 458 460-461 463 466-470 474 476 486 491-493 496 498-500 507 511 514 520-521 525 527-529 531-532 534-535 542 546 548-549 551-552 557-558 560 562 564-566 568 571-572 578-583 586-587 590-591 594 599 602 606 618-619 621 626 629 631-634 643-644 647 651 656-660 664 670 672 677 680 684 687-688 691-695 697 706 709-710 712-714 716-718 721-722 724-725 727-728 730 733 740-741 745 751-752 754 761 765 774 777-779 787 790 792-793 799 801-804 808 810 812 820 822 824 827 831-832 834 836 845 850 858-861 868-869 871-872 875-876 883 887 891 897 900 904 907 910 913 917-920 925-927 929 931-934 938-941 946-947
adult brain	Clontech	ABR011	51 133 810 892
adult brain	GIBCO	AB3001	16-17 19 40 66 92-94 97 124 131 134 163 186 188 208 213 231 268-270 284 288 295 297 299 311 315-325 340 373 387 396 407 429 469 489 495 498-499 533 542 545 562 568 587 589 618-619 643 664 687-688 694-695 730 748 836 876 882 884 902 925-926 948
adult brain	GIBCO	ABD003	2 22-24 29 33 43 45 50-51 66 71 75 77 82 87-88

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
			91-92 95 131 140 157 179 188-192 200 208 213 220 225 247 252 257 261 263-265 277 284 288 295 299 301 315-325 355-356 373 387-389 392 395-396 407 423 431 443-444 450-451 457 459 468 476 489 495 499-500 514 520-522 532-533 541-542 545-546 557-558 562 564 576-577 581-583 588-589 591 595 597 599 601 610 619 631-632 639 643-644 654-655 658-660 664 667 676 682 687-688 693 696 700 704 711 713-714 746 758 765-766 774-775 780 800 802 804 807 810 827 829 834 842 850 854-855 866 870-871 878 892-893 897 899 910 916 920-921 929 931-932 934
adult brain	Invitrogen	ABR014	2 51 65 84 86 134 311 316-325 384 422 445 460 503 525 564 634 651 721 794 804 810 922
adult brain	Invitrogen	ABR015	37 134 263 272 277 294 311 443 467 500 514 582-583 619 651 694 850 871-872 883 888 936
adult brain	Invitrogen	ABR016	19 22 57 134 188 233 271 277 299 373 440 444 459 469 514 640 717 882 890 920
adult brain	Invitrogen	ABT004	1-2 18 28 51 55-57 67 87-88 115 119 137-139 142 163 200 204 213 218 257 263 271 282 288 299 301 311 341 358 370 378 402 407 422-423 427 458 460 463 499 504 534-535 551 557-558 571 586 605-606 610 618 627-628 640 643 680 687 691-692 697 701-702 715 719-721 725 727 753-754 758 771 782 810 827 859 871-872 881 913 920 925-926 938-941 944 946
adult heart	GIBCO	AHR001	1-2 5-6 14-18 20-21 23 28 32 37 41 45 51 53 55-56 62 66 69-70 80-81 85 87 91 97 107 120-121 124 134 140-141 156 163 165-166 172 188-192 195 197-198 200 208 213 216 221 229 231 235 261-265 267 271 276 284 288 302 305 308 311 316-325 328 333-334 337-338 347 368-369 373 376-377 379-380 389 396 420 440 445 453-454 459-460 465 468 478 483-484 489 491-493 495 501 504 507 514 524 529 533 539 541-543 545 549 552-553 564 566 568 574 577 581-583 587 589-591 596 599 602 605 608-609 618-619 623 625 629-632 643 645 647 651 664 672 676 678 683-684 707 714 716-717 732 735 740 743-744 751 754 757 765 775 778 784-786 788 807-808 810 826 828-829 842 850 860 876 878-880 890 894 897 899 902 916 923-927 933 939-941
adult kidney	GIBCO	AKD001	1-2 5-6 13 16-17 19-23 26 28 33 38-39 43 45 48-51 55-57 60 66-67 69-73 79 82-83 87 90 94 96-97 100 103 126 131 134 140 148-149 157 163 166 179 184 186 188-192 200-203 213-216 220-221 224 226-229 232 235 245 252 257 261-263 268-270 272-274 276-277 279 282 288 290 294 299 308 311 316-325 332 335 339-340 358 360-363 373 375 379-380 386 388-389 392 395-396 402 413 421 423-424 428-429 431 436 440 444 450 454 457 459-460 468-469 476 489 492-493 499 504 511 513-514 520-521 524-526 531 533 538-542 544-547 552 564 567-568 574 577-578 582-583 590-591 595-596 598 602 607 610 613 618-619 622 631-632 639-642 644 647 651 654-655 658-659 664 667-669 673 678 680-682 684

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
			687 689 693 696 706-707 712 714-715 717-718 721 729-731 734-736 740 744 748 754 760 771 774 782 784 789 795-807 809-810 819 825 834 836-837 842 850 859-870 872 876 878-879 884 887 890 895 897-899-902 905 910 919-921 925-926 933 936 944
adult kidney	Invitrogen	AKT002	1 14-15 28 30 35-37 53-54 73 88 112 114 129 134 137 140 149 157 166 172 186-188 191-192 203-204 213 235 245 257 262-263 266 268-270 273-274 288-289 297 299 302 310-313 315-325 335 340 358 373 378-381 395 413 423 441 450 453 456 459-460 470 477 491-494 500 513 540 542 545 554 556 564 567 587 591 619 622 627 633 643 668-669 677 684 689 693 701-702 704 714 729-730 754 758 760 777 781 785-786 788-789 807 836-837 840 849-850 872 876 881 890 895 905-906 913 923 925-926 931-933 944
adult liver	Clontech	ALV003	159 179 189-192 201 219 257 349 392 568 664 753 796 887 934
adult liver	Invitrogen	ALV002	5-6 28 35-36 52 54 70 72 86-87 103 112 127 134 140 159 179 188 200-201 213 218-219 225 239-240 257 263 271 275 311 315 367 373 388 392 444 459-460 464 468 497-499 512 527 532 542 545 562 599 605 629 640 657 680 684 687-688 706 713 715 717-718 721 742 754 758 771 791-793 818 829 843 854-855 871 878-879 887 921 933-934
adult lung	GIBCO	ALG001	5-6 16 28 38 51 74 97 122 124 134 140 163 188-192 200 218 221 262-263 268-272 294 311 316-325 379-380 429 463 468 493 511 520-522 537-538 542 545 568-569 595 622 643-644 664 667 711 714 721 730 754 775 850 860 863 879 887 897 925-926 944
adult lung	Invitrogen	LGT002	2 5-8 13 16-17 29-31 35-39 43 46 57 67 72 76 78 81 85 87 90 94 97 100 110 119 122 130-131 134 137 140 146 149 167 172 174 179 188 197-198 201 213 216 218 220-221 223 231 245-246 251-252 256-257 262-263 267-270 277 284 288 296 299 301-302 311 316-325 340 354 373 379-380 388 392 395 400-401 410 413 421 431 436 441-443 445 451 455 457 460 463-464 467 469 475 478 489 491 493 497 499 504 507 514 518-519 524 529 534-535 537 542 545-546 548 552 555 559 568 578 581-583 592 597 602-603 605-607 613 615 619 621-622 636-637 642-643 646-647 654-655 679-681 684 687-689 693 701-702 704 706 711 713 715-716 718 727 732-734 738 748 753-754 757-758 760 762 766 769 774 782 785-786 802 817 829 834 850 853 859-860 866-867 870-871 878-879 887 890 899 902 904 910 917 923 925-926 936-937
adult spleen	Clontech	SPLc01	33 38 57 67 75 87 134 142 163 216 221 229 244 257 304 307 311 316-325 340 355-356 378 441 468 525 538 545 560 564 599 721 754 766 780 794 827 841 850 866
adult spleen	GIBCO	ASP001	2 14-15 20-22 29 38 43 48 51 53-56 65 67 72 74 84 87 131-132 134 137 140 172 188-192 200 212 221 256 263 271 282 308 311 316-325 343

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
			383 389 423 436 441 443 459-460 467 469 495 499-500 505 514 520-522 524 529 537 539 545 552 585 619 631-632 639 643 664 673 707 723 735 742-744 758 771 799 810 817 836 850 878 925-926 934 936
bladder	Invitrogen	BLD001	5-6 8 20-21 28 72 91 122 126 130 166 188 197 200 213-214 225 257 262 315-325 341 409 486 491 572 593 622 650 673 691-692 810 813 861 870 877 883 887 904
bone marrow	Clonetech	BMD007	65 76 84 245 516
bone marrow	Clontech	BMD001	8 13-16 28 38 43 45-48 50-51 57 62-63 65 67 84-85 97 100 104 118 122-124 131 134 140 163 188 214 216 221 224 231 245 252 261-263 268-270 273-274 279 288 290 311 373 378 389-391 395 414 428 431 436 440-441 443 451 455 459-460 465 469-470 475 495 497-498 502 507 514-519 529 537-538 542 546 550 552 556 560-561 563-564 568 576-577 580 587 589 596 601-602 610-613 619-620 626 642-643 647 651 664 666 668 676 678 681-682 684 696 704 706-707 715 727 730 732-735 740 748 753 758 761 764 771 775 780 794 800-801 830 834 836 842 850 863 871-872 878-879 882 884 888 897 900-901 904 910 921 923 929 934 947
bone marrow	Clontech	BMD004	65
bone marrow	GF	BMD002	1-2 5-6 10 13 16-21 27 31 38 42-43 46 57 65-66 76 80 84 87 97 99 110 112 118 131 134 137 140 145 161 163 165 172 195 206 208 221 229 231 237 244 247 252 256 267-270 272 276 278-279 282 284 288 294 301 304 307 311 316-327 333-334 337-338 345-347 352 360-361 368 373 376-378 381 383 388 414 436 441 443 450 452 454-455 457 469-470 483-484 486 490 498 516 519-521 524 530-531 539 542-543 545-546 551 553 555 559 564 571 576-577 580 585 591 594 602 604-605 607-608 610-612 619-621 625-626 629 631-632 639-640 644 650-651 664-665 684 687-688 693 699 703 714 723-724 727 733 735 740 742 745 748 750-752 754-755 777-780 784 787 794-795 802 809 817 824 827 831-832 834 846-847 850-851 854-855 861 867 875 878 883 886 891 894 897 900 902 910 914 919 921 925-926 929 936 939-941 944
cervix	BioChain	CVX001	3-4 14-16 20-23 25 33 42-43 45 48-50 54 57 67 69 75 85 87 91 95-97 107 110 114 124 126-127 131 134 137 140 150 157 163 165 172 185-188 200 204 212-213 216 225 229 245 252 257 261-263 266-270 276 282 288 290 301-302 308 316-325 327 340 363-364 372-373 378 383 388-392 394 396 409 413-414 421 428-429 438-440 443-444 454 456-457 459 463 467 475 486 489 493 495 507 514-515 522 534-537 556 568 572 574 577 582-583 587 594 600 608 610 613 622 626 633 639 643 647-648 651 653 667 680 683 685-686 693 696 703-704 706 711 721 723-725 727 730-731 734-735 742-743 748 754-757 762 771 776 785-786 788 794 800 802 807 809-810 817 827 829 834-835 842 850 857 860 862-863 868

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
			870 873 876-877 879-880 884 887 891 897 904-905 910 916-917 921 925-926 933 937 947
colon	Invitrogen	CLN001	19-21 53 55-57 72 88 133-134 168 213 245 252 311-313 316-325 340 443 459 469 483-484 486 497 515-516 597 606 622 643 667 676 706 718 742-743 753 766 829 833 872 887 902 923 929
diaphragm	BioChain	DIA002	305 311
endothelial cells	Strategene	EDT001	1-2 7-8 14-16 19-22 24 28-29 32-33 41 43 45 51 57 61 74 83 87-88 97 105 112 116-117 131 134 137 140 148 165 172 179 188-192 197-198 208 212-213 220-221 225 229 231 237 246 252 256-258 261-265 268-272 276-277 279 281-282 284 286 288 294 297 299 302 307-308 311-313 326 334-335 340 355-356 358 360-361 364 375 383 386 389 392 403 413 423-424 429 440 443 445 451 453 455-456 459-460 462-463 465-466 468-470 475 491 495 497-499 504 514 520-522 524-526 528 532-536 539-540 546 551-552 554 556 564 566-567 571 574-577 581-583 587 591-593 597-599 601 607 615 618 622 625 633 639 641-644 651 667 677 680 684 691-692 701-702 704 716-717 720-721 726 732-733 735 743-744 754 758 765 785-786 795 802 806 809 819 826 828-830 832 834 836 846-847 850 867 871 877-878 890-891 897 902 907 921 923 925-926 944 946
esophagus	BioChain	ESO002	188
fetal brain	Clontech	FBR001	33 49 51 126 134 197-198 264-265 360-361 413 460 647 810 819 871
fetal brain	Clontech	FBR004	137 156 205 282 284 405 424 480 489 701-702 820 921
fetal brain	Clontech	FBR006	2 9-10 18-19 22 28 30-32 37 39-40 42-43 46-47 49 57 66-67 76 80 83 96 109 112 116-117 120 124 131 133-134 136 142-143 146 152 155 160 162 165 169 173 184 189-198 200-201 205 215-216 238 244 248 254-255 257-258 260-263 272-274 276-277 282 288 293-294 307 309 311 314-328 343 347 351-352 354 357-358 360-361 373-375 378-381 390 392 400-401 403 405 407 410-411 413 420 424 429 445 450 452-453 458 460 463 467-469 472 474 477 479 483-484 491 499 507 520-521 525 527 529 531 533 538 545 551 562 564 566 571 574 579 581-583 587 591 599-600 604 606 611 626 629 631-632 638 643 651 654-655 657-660 672-673 676-677 684 689 693-694 697 699 709 714-715 717 720-721 732-733 735 744 748 751-752 754 761 763 767 772 775 777-779 781 785-786 790 792 802 804 808 810 820 824 826 838-840 850 858-860 864 866 872-873 881 891-892 901-902 904 910-911 913 917-918 920 925-926 933 939-941 946-947
fetal brain	Clontech	FBRS03	316-325 684
fetal brain	GIBCO	HFB001	2 12 16-17 19 23 27-28 32-33 39 41-45 49 87-89 94 97 100 107 112 130-131 134 142 157 163 172 188-192 200 216 224-225 231 237 242 246 252 258 261 263-265 271 273-274 276-277 288 295 299 301 307 311 314-326 328 341 355-356 373 375 387 389 392 395 424-425 431 438 445 450-452 457 459-460 468-469 475 489 491 495

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
			500 504 511 514 520-529 531 533 540-542 545 552 554 557-558 566 576-577 579-584 587 591 596 598-599 606 613 626 631-632 643 651 664 668 673 676 680 693-694 696 703-704 716-717 721 727 735 738 740 744 748 757-758 769 774 778 780-781 810 827-828 830 850 869 871-872 876 878-879 884 890 892 897 899 904 906-907 913 916 918 920 924 928 934 938 946
fetal brain	Invitrogen	FBT002	2-4 20-21 45 51 53 57 88 93 125-126 134 166 184 186 188 200 213 224 263 276-278 307 311 341 373 375 418-419 423 427 432 450 452 459-460 470 492 498-499 507 514 522 534-535 545 550-552 571 577 610 714 721 743 754 795 827 861 866 872-873 887 896 925-926 934 939-941 946 948
fetal heart	Invitrogen	FHR001	2-4 10 13 16-18 29 31-32 37-38 43 46 49 51 53 55-56 67-68 75 80 85 87 97 115 120 137 152 156 160-161 163 168-169 174 178 189-192 196 200 216 220 225 252 262 276-277 282 288 301-302 305 311 315-325 333 343 351 357-358 360-361 368-372 378 424 436 440-441 445 453 460 469 478 483-484 495 520-521 527 533 538 541-543 546 556 564-566 568 576 581 587 594-595 601-602 606 609 612 615 633 638 640 643 653-655 664-665 672-673 677 684 691-693 697 704 707 709 717 735 738 744 746 748-749 751-752 754 761 777 779 781-782 785-786 797-798 820 824 826 829 834 838 841-847 850 875 877-878 893-894 897 901 910 913 925-927 936 946
fetal kidney	Clontech	FKD001	8 14-15 32 43 50 68 96 106 126 131 134 140 186 188 226-228 233 279 282 311 339 428 440 450 456 468 552 618 651 700 726 735 748 751 781 794 797-798 826 878 887 899
fetal kidney	Clontech	FKD002	50 83 96 131 134 143 163 172 193-194 201 203 215 263 273-274 311 316-325 339 360-363 374 376-377 379-380 388 394 400-401 403 407 425 440 451 454 493 525 536-538 540-542 572 580 582-583 587 605-606 621 631-632 647 673 689 706 709 714 726 735 761 774 777 799 809 845-848 858 872 875 878-879 882 895 918 927
fetal kidney	Invitrogen	FKD007	66 214
fetal liver	Clontech	FLV002	52 189-192 219 297 308 335 364 378 427 828
fetal liver	Clontech	FLV004	2 19 28-29 37 39 49 52-53 55-56 62 65-66 76 87 124 134 137 139 142 179 188 195 208 216 219 244 252 263 268-270 272 277 287-288 294 303 305 311 315-325 339 355-356 358 360-361 368 374 378 403 441 454-455 460 477 483-484 497 514 520-521 542 553 582-583 587 591 594 611-613 620-621 638-639 654-655 658-659 681 684 687-688 709 721 730 738 744 752 754 781 793 802 813 818 826 832 836 854-855 876 878 893 897 900 910 924 933 944
fetal liver	Invitrogen	FLV001	2 7 19 28 35-37 47 52 54-56 66 95 134 139 179 188-192 200 213 218 263 272 288 294 305 311 315 349 378-381 388 392 403 426 443 454 459-460 469 496-499 514 527 529 532 534-535 555 586 605 640 644 658-659 673 680 687 698 713 715 720-721 723 726 754 758 778 795 817-818

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
			829 853 861 868 870-871 897 903 933
fetal liver-spleen	Soares	FLS001	1-23 25-39 41-43 45-46 49-65 76 83-84 87-88 91 94 97 100 112 118 122 126 129-130 134 137 140 148 163 168 172 179 186 188-189 191-192 197-198 200-201 213-214 216 221 225-229 231-232 235 242 252-253 256-258 262-271 277 284 287-288 297 299 307-309 311 315-326 330 355-356 360-362 370 373 378-380 388 392 394 396 400-402 413 426 428 436 440-441 443-447 450 454-455 457 459-460 463-465 467-469 475 477-478 489-509 511-514 519-521 525-527 529-535 537 539 542 545 551-552 555-556 559 561 567 569 571 576-577 580-583 586-587 589 591-592 595 598-602 605 607 610 612-613 618-619 623 625-626 631-632 638-640 642-644 646-647 649 651-652 654-655 667 673 676 680-681 683-684 700 703-704 706 711 715-718 720-721 726 732-735 740 742-744 748 754 756-758 763-765 771 774-775 777-782 785-786 790 793-795 797-798 806-808 815 818 824-826 829-830 834-835 837 841 846-847 849-850 856 860-861 866 870-872 876 878 881-883 888 894 897-898 902 905 907 910 919 924-926 929 933 942 947
fetal liver-spleen	Soares	FLS002	1 3-4 11-12 14-17 20-23 26-29 32-34 38 41 43 45-47 49 51-52 55-62 65-67 76 83-85 87-88 90-91 95 97-99 104-105 112 114-117 126 130 133 150 163 165 172 178 186-187 193-194 200-202 208 213 221 225 229 232-235 244-246 248-253 256-257 262-265 267-271 273-274 284 287 299 311 315 326 335 337-338 343 355-356 358 375-378 381 392 394 400-402 414 416 426 428-429 440-442 444-447 453-455 457 459 461 464-465 467 476-477 483-484 489-490 492 495 497-500 504 506-507 509 511-514 519 522 524 526 532-535 537 539-540 542 545 551-552 556 567 569 574 576-577 581 589-590 592 599 601-602 605 607 610 612-613 619-620 625 627 629 631-632 638 640-641 646 648-649 654-655 667 670 683-684 687-688 693 696 700 703-706 713 716-721 726 734-735 740 742 744 748 754 758 771 775 777-778 780-782 785-787 790-792 794-799 801-802 806 808 818 824 829 835 849-850 852-855 857 870-871 876 882-884 886 888 890 894 897-898 900 902-903 907 919 921 923-926 929-930 933 938-942
fetal liver-spleen	Soares	FLS003	30 34 67 85 88 99 130 172 188-189 191-192 213 229 231 257 311 315-325 329 331 335 362 391 394 400-402 423 441 455 457 461 476 498 500 511 523 531-532 537 542 576 587 592 612-613 625 649 665-666 703 719 731 733 740 744 771 775 777-778 787 797-798 819 824 826 850 854-856 861 863 870 879 884 897 923 931-933 947
fetal lung	Clontech	FLG001	65 179 213 223 340 360-361 491 564 577 591 627 646 650 712 715 744 758 939-941
fetal lung	Clontech	FLG004	388 921
fetal lung	Invitrogen	FLG003	49 54 75 97 137 148 152 188 197-199 213-214 225 240 256 288 316-325 369 378 392 423 429 464 496 526 580-581 586 591 693 706 726-727

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
			766 878 913 925-926 939-941
fetal muscle	Invitrogen	FMS001	28 65 115 121 126 134 137 156 168 172-173 181 213-214 225 263 297 305 340-341 360-361 440 459 516 534-535 563 564 586 606 609-610 623 650 676 683 754 766 853 871 886 894 930 934-935 948
fetal muscle	Invitrogen	FMS002	19-21 41 49 51 53 57 75 96 101 103 112 134 136 156 171 184 188 191-192 212 216 250 262 267 276 305 311 342 348-350 355-356 360-361 374 392 403 411 415 423 425 457 469 491 495 499 508 515 517 534-536 543 546 564 566 576 580 582-583 587 594 599 609 611 615 618 623 644 647 658-659 664 668-669 677 683 691-692 696 703 735 743 754 766 788 802 817 826 828 850 877-879 894 910 925-926 935
fetal skin	Invitrogen	FSK001	3-7 18 24 27 29 35-37 51 53 55-56 66-67 76 90 95 97 122 126 134 136-137 166-167 181 188-192 208 213-214 224-225 245 250 252 257 260 262 268-271 273-274 282 284 297 302 312-313 315-326 341 367 373 375 378 383 387-388 390 394 423 429 440-441 450 454-455 457 459 463-464 470 472 475 486 489 492 495 498-500 511 514 524 527 530-532 534-535 541-542 545 550 552 555-556 565 571 586 589 591-592 602 604 606 610 613 618 622-623 626 631-632 640 651 654-655 672-673 685-686 693 701-702 704 706 717-718 720 723 727-728 744 754 762 764-766 768-771 795 809 814 821 824 827 843 853 868 870-874 887-888 890 897 902 907 925-926 928 930 933-934 939-941 944-945 947
fetal skin	Invitrogen	FSK002	2 5-6 19 29 34 51 57 59 88 97 101 124 131 134 143 163 166 172 189 191-192 196 212 216 222-223 231 250 257 263 268-272 282 284 287-288 294 297 299 302 304 310-311 316-325 328 333 340 352 360-361 365-367 372 379-380 388 390 400-401 403 410-411 440 449-450 454 457 463 470 478 491 495 500 505 515 520-521 524 532 534-535 541 555 560 562 564 572 576 581 592 595 599 611 622 626 630 636 640 642 650 664 677 683 691-693 696 699 701-702 708-709 715 721 723 728 735 744 747-748 750 754 766 779 782 799 803 807 813 820 824 826 834 846-847 867-868 872-874 878-879 890-891 897 901 904 907 910 912 916 918 925-926 933 944-945
fetal spleen	BioChain	FSP001	311 748
infant brain	Soares	IB2002	2-4 12 14-15 20-21 23-24 27 29 31-32 39 41 46 48-49 51 53 55-56 66 75-76 86-88 93 95 101 105 108-109 116-117 125 127 129 131 136 145 163 166 170 180-181 186 188-189 191-192 200-201 207-208 212-214 216 220 224 229 231 245 247 252 257 259 264-265 267 271 279 282 288 293 295 299-300 311 314-326 337-338 340 349 367 373 375 388 390-393 396 402 405 407 418-421 424 428-429 431 433 436 450 452-453 457 459 463 468 489 495 498-500 507 511 522-524 526 528-530 532 541-542 545-546 552 557-558 562 564-566 571 577-583 587 589-591 599-601 606 608 613-614 619 631-632 647 654-655 658-

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
			659 667 676 684 691-693 696 700 704 711 718 721 723 725 740-741 743-744 748 754 775 777-778 780 788 792-793 795 802 805 808 819 826-829 834 836 838 861 863 869-870 875 879 881-882 884 887 890-891 893 897 902 920-921 925-926 934 938-941 946
infant brain	Soares	IB2003	2 27 37 39 43 48-49 51 53 85-87 97 106 113 124 126-127 131 142 166 170 188 200-201 208 214-215 220 224 226-228 231 251 257 263 267 271-272 279 288 293 299 311 314-326 337-339 349 360-361 367 386 392 397 400-402 407 410 418-419 424-425 427 429 452 454 460 475 489 495 497-500 507 522-523 525 529 532 539 542 545-546 551-552 557-558 564-565 578 582-583 585 591 601 606 625-626 631-633 643-644 673 690-693 701-702 706 711 721 723 734 740-741 743-744 748 751 754 761 778 788 795 802 808 819 826-827 829 837 843 869-871 875 878-880 884 896-897 902 920 933-934 946
leukocytes	Clontech	LUC003	12 14-15 18 32 111 134 137 172 221 277 280 311 316-325 436 454 467 549 552 568 585 603 643 691-692 698-699 734 744 751 754 784 797-798 861 897 916 923
leukocytes	GIBCO	LUC001	2-4 7-8 13-17 20-23 31-33 38 43 48-49 51 53-57 63 66-68 74-78 85-88 93 97 122 124 129 131-132 134 137 140 163 166 168 171-172 175 188-192 197 200 208-213 216 221 223 231 236 242 252 257-258 261-263 268-270 272 277 279 287-288 294 307 311 314 316-326 329 337-339 341 373-374 376-377 381 388-392 396 400-401 413-414 423 436 441 450 454-455 459 463 465 467 489 491-493 495 498-499 504-505 507 514 518 520-522 524 526 529 531 533 536-537 539-540 545 552-554 556 568 571 577 580 585 589-590 596 599 602 605 607 610 612-613 615 618-619 621-622 625 638 640 642-644 664 667 677-678 684 690-693 696 700 703-704 707 713 715-718 721 727 734-735 738 740-746 748 753-754 758 775 778 780 789 794 797-798 801-802 815 817 825 827 829 834 836 846-847 850 859-861 863-864 866-867 871-872 878 884 886-888 891 896-897 902 904 910 913 916 921 923-926 929-932 936 943
lung	Strategene	LFB001	2-4 22 28 32-33 47 51 79 120 129 134 140 163 172 188 208 220-221 231 252 257 263 276-277 284 307 375 378-380 396 423 428 440 450 459 463 486 491 493 495 499 539 571 591 601 607 613 615 618 625 639 651 684 716-717 721 727 735 748 782 828 850 870-871
lymph node	Clontech	ALN001	43 98 131 140 163 188 221 245 277 299 311 491 515 546 564 593 603 610 615 630 682 694 707 717 800 831 850 878 880 936 939-941 947
lymphocyte	ATCC	LPC001	2 16 19-21 25 31-32 49 53 55-56 63 67 85-87 90 97 120 122 137 140 163 165 168 172 188 197-198 215-216 221 229 231 236-238 248 252 256 272-274 283-284 288 294 299 316-326 343 368 374 378 395 423 431 454-455 467 469 476 478 491 495 498 505 512 515 517-518 520-522 524

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
			526 529 531-532 537 539 542 545-546 551 556 571 577 580 589-590 592 596 601 622 631-632 640 642 654-655 664 666 668-669 673 684 703 708 716 718 721 723 727 733 735 743-744 754 758 765 771 775 777-779 783-786 797-799 810 816-817 828 834 845 859 861 863 870 878-879 881 884 887 890 897 904 907 910 912-913 918 923 929 939-942 945-947
macrophage	Invitrogen	HMP001	49 97 208 252 301 306 311 316-325 337-338 345-346 416 512 522 572 670 716 743 785-786 802 888 919 923
mammary gland	Invitrogen	MMG001	1-8 14-18 20-21 25 28-29 37 39 43 49 51-57 60 66-67 72 75-76 87 95 97 103-104 106 112 115 119 122 127 130 134 137 139 142 150 166-168 172 175 184 186 188-189 191-192 200 213-214 222-224 226-229 240 252 257-259 263 267 271 276 278 282 287-288 299 301-302 305 307-308 311-313 316-325 327-328 332 340-341 358 360-362 369 373 378 381 383 388 390-392 397 402-403 409 415-416 423 425 428-429 433 436 444 454-456 459-460 464 467 469 481 483-484 486 493 495 498-499 515 524-525 529-530 532-537 541-542 545 551-552 562 582-583 586-587 593 599-600 602 604-605 610 618-619 622 625-627 634 644 646-647 652 654-655 662 673 676 680 684 687-688 691-692 701-703 715 717 721 723 726 735 743 751 754 758 765-766 771 777-778 789 803 805 807 809-811 821 827 829 850 860 887-888 892 896 898 901-902 905 911 913 917 925-926 930 936 939-942
melanoma	Clontech	MEL004	3-4 16 20-21 43 46 48 97 103 147 163 188 191 213 216 221 231 241 245-246 260 262-263 316-325 381 407 431 504 525 527 542 556 568 577 589 596 607 613 676 693 714 735 737-739 744 758-760 775 822 850 863 878 887 897
neuron	Strategene	NTD001	2 16 32 51 66 88 97 124 130 134 137 172 188-189 191-192 231 252 257 260 277 291 373 424 431 454 460 489 495 523 525 582-583 591 631-632 643 649 670 695 725-726 735 765 789 797-798 837 850 878 884 888 890 913 929 946
neuron	Strategene	NTR001	2 5-6 20-22 136-137 188-194 197-198 224 311 375 381 410 457 462 475 495 531 546 548 552 599 618 678 743 752 819 828 890 895 897 930 934 938 944 946
neuronal cells	Strategene	NTU001	2 5-6 20-21 55-56 87 137 188-192 197-198 215-216 260 287 291 310-311 316-325 365 375 423 457 459 470 499 532 542 564 576 598-599 623 643 651-652 673 721 726 743 745 752 754 765 780 787 789 822 829 870 875 888 896 917 919 929
ovary	Invitrogen	AOV001	2-7 10 13 18-22 25 27-28 30-31 33 38-39 41-43 45 48 50-51 53-56 62 66-67 69 72 74-75 80 83 85-87 93 95 99-101 107 112-115 120 124-126 129 131 134-137 140 142-143 147-148 162-163 172 178 188 191-192 200 204 208 212-213 220-221 225 229 231 235 237 246-247 252-253 258 261-262 264-265 267-272 276-279 282 287-288 290 293-294 299 307-308 311 316-327 332 337-

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
			338 340-342 349 360-362 373-374 379-381 386 388-389 393 396 399 403 413-414 423 425 427 429 431 441 444-445 450 452 454-455 457 459-460 462 467-470 475 477 483-484 489 491-493 495-496 500 504-505 507 515 518-519 522-523 527-529 531 533 537-540 542 545-546 548 551-552 555 564 568 570-571 577 579 581-583 589-596 599-600 605 610 613 615-616 619 623 625-627 630-636 639 641-644 647 649 651 654-655 664 668-669 672 676 678 680-682 684 687 694 701-704 706-707 715-717 721-722 727 729 731-732 734-735 738 740 743 748-749 753-754 758 764 771 775 777-782 784 800 802 807 821-822 824 828 834 836-837 842 846-847 850 860-863 866 870-871 876-880 882 884 887 890-891 897 899 901 906 910 913 920-921 923-924 933 939-941 944 947
pituitary gland	Clontech	PIT004	41-42 83 85 97 134 193-194 204 208 213 224 257-258 263-265 285-286 308 311 360-361 413 443 445 491 514 529 532 639 644 647 682 701-702 716 781 822 829 836 850 933 939-941 947
placenta	Clontech	PLA003	16 31 34 49 66 80 87 97 101-102 134 158 165 172 179 184 188 197-198 209-210 218 220 229 235 249 256 267-270 277 287-288 302 307 332 360-361 365 388 394 414 441 444 454 457 460 493 498-500 505-506 509 529 531-532 550 559-560 564 572 587 601 625 630-632 638 672 682-684 689 706 708 726 733 735 744 754 761 784-786 793 863 875 897 924 929 937
placenta	Invitrogen	APL001	34 68 102 263 444 493 520-521 534-535 689 706 754 797-798
placenta	Invitrogen	APL002	2 14-15 43 55-56 66-67 134 184 213 221 229 252 257 263 277 287 394 443 529 532 618 622 684 742 754 810 829 883 902
prostate	Clontech	PRT001	7-8 51 85 87 97 100 122 134 139 214 216 221 231 257 271 276 335 337-338 392 400-401 431 440 459 477 530 534-535 546 556 582-583 599 622 631-632 639 651 663-664 673 683 707 715 735 740 765 773-774 777 810 823 897 909 919 934 939-941 947
rectum	Invitrogen	REC001	18 54 66 134 137 169 188 200 213 225 251 263 288 311-313 316-325 340 388 423 429 441 454 459 514 532 542 610 626 646 651 657 715 719 723 728 735 740 758 766 785-786 823 829 833 836 886 942
saliva gland	Clontech	SALS03	460
salivary gland	Clontech	SAL001	31 49 78 95 134 136-137 143 176 188 208 223 244 268-270 284 308 311 316-325 388-389 391 436 441 459 476 514-515 520-521 532 543 568 589 596 610 619 684 691-692 713 718 727 736 754 777 824 836 864 867 878 883 897 901-902 916-917 933 938-941
skeletal muscle	Clontech	SKM001	42 98 156 163 191-192 200 261 305 311 395 415 462 468 504 531 543 566 582-583 585 594 680 740 853 875 927 933 935
skeletal muscle	Clontech	SKM002	850
skin fibroblast	ATCC	SFB001	379-380 850
skin fibroblast	ATCC	SFB002	742 850

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
skin fibroblast	ATCC	SFB003	87
small intestine	Clontech	SIN001	27-29 31 38 40 46 48 51 54 57-58 62 65 67 75 77 85 97 110 112 116-117 119 131-132 134 137 140 161 163 166 168 177 188 197-198 208 213 220 224 229 246 257 261-262 264-265 276-277 288 295 297 299 311 316-326 328-330 337-338 340 360-361 373 375 382 390-391 410 413 428-429 436 438 440 453-454 459 468 476-477 497 507 511 522 531 536 538 542 545-546 548 552 556 564 570-571 576 580-581 586-587 591 596 599 605 610 613 619 625-626 643-644 651-653 664-666 668-670 677 680 684 693 700-702 706-707 713-715 723-724 729-730 735 740 746 748 753-754 757-758 764 777-778 784-786 818 822 824 826-829 833-837 842 862-863 865-867 877-878 886 897 900-902 906 913 916 921 925-926 936 939-941
spinal cord	Clontech	SPC001	18 23 33 37 42 51 67 87 92 94 97 100 140 162 184 188 191-192 208 213 220 231 248 262 268-271 273-274 282 287-288 290 307 311 316-325 358 364 376-377 383 387 389-390 402 412 422 444 455 476 483-484 489 504 522 534-535 556 562 587-588 591 597 603-604 618-619 643 651 667-670 677 693 703-704 717-718 727 746 757 773 808 810 827 834-835 837 850 871-872 875 904 910 931-932 939-941
stomach	Clontech	STO001	18 65 88 163 188 208 213 261 272 277 286 294 336 373 396 412 459 514 553 602 610 647 651-652 671 673 714 774 790 831 833 842 850 876
testis	GIBCO	ATS001	1 3-4 14-16 28 31 45-46 66 85 90 95 97 103 112 128-130 134 140 163 166 188 191-192 199-200 213 226-228 261-265 267-271 284 302 311 316-325 327 379-380 391 413 421 428 444 454 457 459-460 467 491 493 495 500 505 519 525 529 532 534-535 545 552 556 566 568 575 596 599 613 616-617 647 649 651 680 684 703 707 716 719 721 727 734 738 740 744 748 758 765-766 774 777 782 802 810 817 827-828 834 842 846-847 850 862-863 871-872 878 880 892 901 916-917 921
thalamus	Clontech	THA002	2 87 96 103 106 189-192 208 252 258 295 308 311 367 376-377 383-384 445 455 459-460 498 529 587 598 602 629 654-655 705-706 715 717 723 754 775 810 817 822 864 867 881 892 927 930
thymus	Clonetech	THM001	3-4 8 18 28 54 57 63 65 68 84 97 100 116-117 122 134 142 151 169 171-172 188 195 197-198 201 213 221 237 245 261 287 311 316-325 360-361 376-377 423 441 444 459 489 491-493 495 498 504 507 514 527 532 534-536 539 553 556 568 571-572 590 595-596 599 610 618 622 631-632 643 647 651 654-655 664 687-688 691-693 703 715 721 733-735 748 760 762 765 781 794 799 802 831 834 836 842 850 860-861 863 871 878 885 896-897 903 910 923 925-926 928 939-941
thymus	Clontech	THMc02	2-4 17 20-22 37-38 42-43 46 63 65-68 76 88 95 103 118 120 124 134 137 140-141 143 163 165

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
			171 179 182 189-194 198 200 212-215 221 226-228 231 244 257 262 266 276-277 287-288 297 299 307 316-325 341 352 358 360-361 375 376-377 379-381 389 391 394-396 403 410-411 436 440 445 450 459 463-464 469 478 491 495 500 507 511 519-521 530 532 539 542 550 555 560 563 576 581 587 595 601 610-611 613-614 618 622 625-626 631-632 638 642-644 657 664 667 670 673 680 683 687 691-693 699 715-716 721 740 743-744 747-754 761 763-765 771 777 780-781 784-787 790 794 805 811 820 826 831 834 841 845 861 867-868 878 881 883 891 893-894 896-897 902 904 910 912-914 918 923 936-941 946-947
thyroid gland	Clontech	THR001	1-2 18-21 27 32 38 42 46 49 51 53-56 66 72 77-78 87-88 97 115 119 124 130-131 134 136 152 163-165 172 183 188-192 202 212-213 216 221 224 229 235 241-243 252 257-258 261 263-265 267 277 279 297 301 305 308 311 316-325 327 357 363 373 376-377 381 383 389 397 400-401 410 413-414 427-428 443-444 446-447 457 459 463 467-469 475 482 489 495 499-500 504 509 513 519-522 526 529 533 537-538 542 545-546 548 556 564 567-568 582-583 589 592 599 605 608 611 621 623 630 642-644 648 651-652 654-655 664 672-676 684-686 691-694 700 706-708 713 717-718 721-722 725 729 731 734-735 740 748 753-754 760 764 766 771 774 777 781 792 797-800 802 805 826 828-829 834 842 850 861 863 868 876 879 897 899 901 910 913 929 937 939-941
trachea	Clontech	TRC001	20-21 38 112 161 163 188 263 267 327 413 420 457 459-460 471 514 540-541 552 572 574 622 639 654-655 676-677 691-692 707 725 743 748 765 777-778 862 868 897 905 908 944
umbilical cord	BioChain	FUC001	1-2 29 32 46 67 83 87 94 134 136 140 148 160 163 166 172 181 186-192 197-198 208 213 216 225-231 237 252 261-265 267-270 279 282 288 295 302 308 311 316-326 339-340 365 376-377 379-380 384 392-397 421 423 428 433 440 445 452 459 461 463-464 470 472 489 491 495 497 500 507 517-518 522 525-526 528 534-535 540 545-546 556-558 564 566 568 571-572 577 592 599 601 605 610 618 623 644 651 661 668-669 673 678 680 685-686 696 706 709 718 735-736 748 754 769 772-777 782 792 797-799 802 807 809 815 817 824 850 854-855 870 876 881 888 891 897 899 901 913 921 928 930-932
uterus	Clontech	UTR001	51 67 126 130 133 140 188-192 229 267 329 373 440 491 514 599 685-686 693 713 716-717 735 897 905 911 939-941
young liver	GIBCO	ALV001	3-4 17 20-21 32 43 55-56 70 100 134 137 163 172 174 179 186 188-192 200 213 216 219 221 229 232 252 275 301 311 315-325 378 381 392 441 459-460 497 499-500 514 524 526 533 539 550 568 571 588-589 595 619 622 631-632 642 658-659 664 677 680 693 700 707 713 719 743 754 757-758 766 807 834 863 867 876 884 887

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
			904 907

TABLE 2

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
949	AAM25384	Homo sapiens	HYSE- Human protein sequence SEQ ID NO:899.	644	99
949	AAY27581	Homo sapiens	HUMA- Human secreted protein encoded by gene No. 15.	644	99
949	gi13785618	Mus musculus	sideroflexin 4	396	60
950	gi1276912	Homo sapiens	Human putative ubiquitin C-terminal hydrolase (UHX1) mRNA, complete cds.	3719	100
950	gi12653165	Homo sapiens	ubiquitin specific protease 11, clone MGC:8620 IMAGE:2961383, mRNA, complete cds.	3709	99
950	gi13529494	Mus musculus	Similar to ubiquitin specific protease 11	3167	83
951	AAY11696	Homo sapiens	MITU LckSH3 domain-combining protein.	4131	99
951	AAG78648	Homo sapiens	SHAN- Human Ra1BPI related protein 82.	3875	99
951	gi13625166	Homo sapiens	RALBP1 mRNA, complete cds.	3875	99
952	AAY11696	Homo sapiens	MITU LckSH3 domain-combining protein.	3953	96
952	AAG78648	Homo sapiens	SHAN- Human Ra1BPI related protein 82.	3697	96
952	gi13625166	Homo sapiens	RALBP1 mRNA, complete cds.	3697	96
953	gi10437191	Homo sapiens	cDNA: FLJ21146 fis, clone CAS09305.	2190	94
953	gi12843222	Mus musculus	putative	1672	77
953	gi15778893	Homo sapiens	Similar to hypothetical protein FLJ20967, clone MGC:11140 IMAGE:3837082, mRNA, complete cds.	1328	99
954	gi10437191	Homo sapiens	cDNA: FLJ21146 fis, clone CAS09305.	2359	100
954	gi12843222	Mus musculus	putative	1643	72
954	gi15778893	Homo sapiens	Similar to hypothetical protein FLJ20967, clone MGC:11140 IMAGE:3837082, mRNA, complete cds.	1409	99
955	gi15825377	Mus musculus	NIMA-related kinase 8	2009	89
955	gi15825379	Danio rerio	NIMA-related kinase 8	1439	70
955	AAO019	Homo sapiens	HYSE- Human polypeptide SEQ	548	85

SEQ ID	Hit ID	Species	Description	S score	Percent identity
	74		ID NO 15866.		
956	AAW88660	Homo sapiens	HUMA- Secreted protein encoded by gene 127 clone HSUBW09.	175	97
956	AAO00187	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 14079.	70	55
956	gi13815429	Sulfolobus solfataricus	Sugar transport related protein	70	40
957	AAB93966	Homo sapiens	HELI- Human protein sequence SEQ ID NO:14027.	1133	100
957	gi10433835	Homo sapiens	cDNA FLJ12377 fis, clone MAMMA1002524, weakly similar to HYPOTHETICAL 117.8 KD PROTEIN IN STE2-FRS2 INTERGENIC REGION.	1133	100
957	AAO04381	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 18273.	594	100
958	AAB95297	Homo sapiens	HELI- Human protein sequence SEQ ID NO:17525.	885	100
958	gi10434941	Homo sapiens	cDNA FLJ13087 fis, clone NT2RP3002099.	885	100
958	gi16740566	Homo sapiens	Similar to hypothetical protein FLJ13087, clone MGC:15009 IMAGE:3536735, mRNA, complete cds.	807	95
959	AAY27676	Homo sapiens	HUMA- Human secreted protein encoded by gene No. 110.	474	100
959	gi5295951	Oryza sativa	Similar to Herpesvirus papio BRRF2 homolog gene, partial cds.(U23857)	69	41
960	AAG89262	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 382.	352	98
960	AAY30721	Homo sapiens	HUMA- Amino acid sequence of a human secreted protein.	352	98
960	AAB23615	Homo sapiens	ALPH- Human secreted protein SEQ ID NO: 30.	343	97
961	AAY72605	Homo sapiens	INCY- Human Electron Transfer Protein, ETRN-3.	579	100
961	AAO11627	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 25519.	579	100
961	AAG03941	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 8022.	570	98
962	gi14603455	Homo sapiens	ubiquitin-conjugating enzyme E2A (RAD6 homolog), clone MGC:20175 IMAGE:3051041, mRNA, complete cds.	599	79
962	gi4883773	Gallus gallus	ubiquitin-conjugating enzyme	599	79
962	gi14485244	Mus musculus	ubiquitin-conjugating enzyme HR6A	599	79
963	gi14603455	Homo sapiens	ubiquitin-conjugating enzyme E2A (RAD6 homolog), clone MGC:20175 IMAGE:3051041, mRNA, complete cds.	699	90
963	gi4883773	Gallus gallus	ubiquitin-conjugating enzyme	699	90
963	gi144852	Mus musculus	ubiquitin-conjugating enzyme	699	90

SEQ ID	Hit ID	Species	Description	S score	Percent identity
	44		HR6A		
964	gi168770 66	Homo sapiens	clone MGC:24447 IMAGE:4077762, mRNA, complete cds.	362	100
964	gi168770 59	Homo sapiens	clone MGC:24437 IMAGE:4075637, mRNA, complete cds.	362	100
964	AAY949 59	Homo sapiens	GEMY Human secreted protein clone mc300_1 protein sequence SEQ ID NO:124.	204	97
965	AAB929 93	Homo sapiens	HELI- Human protein sequence SEQ ID NO:11723.	2879	97
965	AAG813 64	Homo sapiens	ZYMO Human AFP protein sequence SEQ ID NO:246.	2879	97
965	gi140423 80	Homo sapiens	cDNA FLJ14690 fis, clone NT2RP2005270.	2879	97
966	AAB957 69	Homo sapiens	HELI- Human protein sequence SEQ ID NO:18703.	2841	99
966	gi104366 07	Homo sapiens	cDNA FLJ14207 fis, clone NT2RP3003185, weakly similar to TROPOMYOSIN 1, FUSION PROTEIN 33.	2841	99
966	gi128331 93	Mus musculus	putative	2375	85
967	AAM254 13	Homo sapiens	HYSE- Human protein sequence SEQ ID NO:928.	799	100
967	AAW678 63	Homo sapiens	HUMA- Human secreted protein encoded by gene 57 clone HFEBF41.	551	98
967	gi135438 11	Mus musculus	Unknown (protein for IMAGE:3591061)	95	33
968	gi104379 60	Homo sapiens	cDNA: FLJ21792 fis, clone HEP00441.	5865	99
968	AAE0618 6	Homo sapiens	HUMA- Human gene 58 encoded secreted protein fragment, SEQ ID NO:248.	3088	98
968	AAE0609 8	Homo sapiens	HUMA- Human gene 58 encoded secreted protein HSLCX03, SEQ ID NO:160.	3088	98
969	gi126980 79	Homo sapiens	mRNA for KIAA1767 protein, partial cds.	4441	98
969	AAM255 78	Homo sapiens	HYSE- Human protein sequence SEQ ID NO:1093.	3898	98
969	AAE0618 6	Homo sapiens	HUMA- Human gene 58 encoded secreted protein fragment, SEQ ID NO:248.	3464	98
970	AAY483 59	Homo sapiens	META- Human prostate cancer-associated protein 56.	403	98
970	gi152159 66	Homo sapiens	DL8Q12 gene for hypothetical protein, exons 1-2.	92	53
970	AAR992 56	Homo sapiens	UYAR- Natural killer lytic associated protein.	75	37
971	gi656182 7	Mus musculus	Kif21a	5684	76
971	gi656182 9	Mus musculus	Kif21b	4944	60
971	gi126979	Homo sapiens	mRNA for KIAA1708 protein,	4656	99

SEQ ID	Hit ID	Species	Description	S score	Percent identity
	61		partial cds.		
972	AAW750 79	Homo sapiens	HUMA- Human secreted protein encoded by gene 23 clone HBMCT32.	148	100
973	AAY359 21	Homo sapiens	GEST Extended human secreted protein sequence, SEQ ID NO. 170.	548	99
973	AAM253 86	Homo sapiens	HYSE- Human protein sequence SEQ ID NO:901.	494	96
973	AAY359 23	Homo sapiens	GEST Extended human secreted protein sequence, SEQ ID NO. 172.	494	96
974	AAY275 87	Homo sapiens	HUMA- Human secreted protein encoded by gene No. 21.	448	100
974	gi128025 61	Bovine herpesvirus 4	unknown	74	42
975	AAU162 97	Homo sapiens	HUMA- Human novel secreted protein, Seq ID 1250.	2420	98
975	AAB944 86	Homo sapiens	HELI- Human protein sequence SEQ ID NO:15170.	1761	74
975	AAM940 18	Homo sapiens	HELI- Human stomach cancer expressed polypeptide SEQ ID NO 106.	1761	74
976	AAM412 64	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 6195.	903	99
976	AAM394 78	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 2623.	903	99
976	AAB437 71	Homo sapiens	HUMA- Human cancer associated protein sequence SEQ ID NO:1216.	903	99
977	gi117616 11	Homo sapiens	kinesin-like protein RBKIN1 (RBKIN) mRNA, complete cds, alternatively spliced.	9290	99
977	gi117616 13	Homo sapiens	kinesin-like protein RBKIN2 (RBKIN) mRNA, complete cds, alternatively spliced.	9055	98
977	gi120540 30	Homo sapiens	mRNA for KINESIN-13A1 (KIN13A gene).	8955	97
978	gi759580 2	Mus musculus	ELKL motif kinase 2 short form	188	48
978	gi759580 0	Mus musculus	ELKL motif kinase 2 long form	188	48
978	AAM939 56	Homo sapiens	HELI- Human polypeptide, SEQ ID NO: 4158.	187	48
979	gi161984 56	Homo sapiens	Similar to RIKEN cDNA 0610040E02 gene, clone MGC:17973 IMAGE:3919892, mRNA, complete cds.	1050	100
979	gi167406 89	Mus musculus	RIKEN cDNA 0610040E02 gene	891	76
979	gi128413 15	Mus musculus	putative	891	76
980	gi147149 27	Homo sapiens	amino acid transporter system A1, clone MGC:17722 IMAGE:3871101, mRNA, complete cds.	2466	100
980	gi116407	Homo sapiens	amino acid transporter system	2466	100

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
	43		A1 mRNA, complete cds.		
980	AAB935 56	Homo sapiens	HELI- Human protein sequence SEQ ID NO:12942.	2459	99
981	AAW750 90	Homo sapiens	HUMA- Human secreted protein encoded by gene 34 clone HTEGA81.	507	100
981	AAW751 52	Homo sapiens	HUMA- Human secreted protein encoded by gene 34 clone HKMLK44.	507	100
981	AAW751 51	Homo sapiens	HUMA- Human secreted protein encoded by gene 34 clone HTEGA81.	507	100
982	AAB947 54	Homo sapiens	HELI- Human protein sequence SEQ ID NO:15811.	4658	99
982	gi140428 59	Homo sapiens	cDNA FLJ14964 fis, clone PLACE4000581, moderately similar to FIBROPELLIN I PRECURSOR.	4658	99
982	gi111771 64	Mus musculus	polydom protein	3880	81
983	AAB652 78	Homo sapiens	GETH Human PRO1185 (UNQ599) protein sequence SEQ ID NO:401.	993	100
983	AAM253 16	Homo sapiens	HYSE- Human protein sequence SEQ ID NO:831.	993	100
983	AAM238 05	Homo sapiens	HYSE- Human EST encoded protein SEQ ID NO: 1330.	993	100
984	AAY359 96	Homo sapiens	GEST Extended human secreted protein sequence, SEQ ID NO. 381.	589	66
984	AAB652 78	Homo sapiens	GETH Human PRO1185 (UNQ599) protein sequence SEQ ID NO:401.	567	65
984	AAM253 16	Homo sapiens	HYSE- Human protein sequence SEQ ID NO:831.	567	65
985	gi168770 39	Homo sapiens	hypothetical protein FLJ22688, clone MGC:2438 IMAGE:2819805, mRNA, complete cds.	1952	91
985	gi133252 53	Homo sapiens	Similar to hypothetical protein FLJ22688, clone MGC:4098 IMAGE:2819805, mRNA, complete cds.	1952	91
985	gi104391 77	Homo sapiens	cDNA: FLJ22688 fis, clone HSI11003.	1695	89
986	gi512469	Homo sapiens	H.sapiens HLA-DMA gene.	1285	92
986	gi218187 6	Homo sapiens	Human DNA sequence from clone XX-O27 on chromosome 6. Contains the BRD2 gene encoding bromodomain-containing 2 protein, the HLA-DMA gene encoding major histocompatibility complex class II DM alpha, two CpG islands, ESTs, STSs and GSSs, complete sequence.	1285	92
986	gi150303	Homo sapiens	clone MGC:13532	1285	92

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
	36		IMAGE:4245221, mRNA, complete cds.		
987	AAB42529	Homo sapiens	CURA- Human ORFX ORF2293 polypeptide sequence SEQ ID NO:4586.	2683	99
987	gi12697893	Homo sapiens	mRNA for KIAA1674 protein, partial cds.	2683	99
987	AAM90536	Homo sapiens	HUMA- Human immune/haematopoietic antigen SEQ ID NO:18129.	1512	98
988	gi2224539	Homo sapiens	Human mRNA for KIAA0299 gene, partial cds.	9903	99
988	AAY16588	Homo sapiens	RHON A protein that interacts with presenilins.	4733	97
988	gi7301710	Drosophila melanogaster	CG11754 gene product	3074	43
989	AAW74887	Homo sapiens	HUMA- Human secreted protein encoded by gene 160 clone HCELB21.	203	100
989	AAM24401	Homo sapiens	HYSE- Human EST encoded protein SEQ ID NO: 1926.	183	85
989	gi9294003	Arabidopsis thaliana	cytochrome P450-like protein	70	39
990	gi11990420	Homo sapiens	mRNA for MOP-3, complete cds.	4359	93
990	AAB93229	Homo sapiens	HELI- Human protein sequence SEQ ID NO:12223.	3950	99
990	gi7023114	Homo sapiens	cDNA FLJ10833 fis, clone NT2RP4001206, moderately similar to Drosophila melanogaster strawberry notch mRNA.	3950	99
991	gi4337105	Homo sapiens	MSH55 gene, partial cds; and CLIC1, DDAH, G6b, G6c, G5b, G6d, G6e, G6f, BAT5, G5b, CSK2B, BAT4, G4, Apo M, BAT3, BAT2, AIF-1, 1C7, LST-1, LTB, TNF, and LTA genes, complete cds.	668	100
991	gi29969	Homo sapiens	Human gene for casein kinase II subunit beta (EC 2.7.1.37).	668	100
991	gi29967	Homo sapiens	Human mRNA for phosvitin/casein kinase type II beta subunit (EC 2.7.1.37).	668	100
992	AAY10840	Homo sapiens	HUMA- Amino acid sequence of a human secreted protein.	349	100
993	AAM25927	Homo sapiens	HYSE- Human protein sequence SEQ ID NO:1442.	824	100
993	AAY73325	Homo sapiens	INCY- HTRM clone 001106 protein sequence.	820	99
993	AAG03870	Homo sapiens	GEST Human secreted protein SEQ ID NO: 7951.	819	99
994	AAB94043	Homo sapiens	HELI- Human protein sequence SEQ ID NO:14201.	5698	99
994	gi10433976	Homo sapiens	cDNA FLJ12471 fis, clone NT2RM1000894, highly similar to DNA-DIRECTED RNA	5698	99

SEQ ID	Hit ID	Species	Description	S score	Percent identity
			POLYMERASE I 135 KD POLYPEPTIDE (EC 2.7.7.6).		
994	gi1621594	Mus musculus	second largest subunit of RNA polymerase I	5095	84
995	AAU15880	Homo sapiens	HUMA- Human novel secreted protein, Seq ID 833.	1091	100
995	ABB03345	Homo sapiens	HUMA- Human musculoskeletal system related polypeptide SEQ ID NO 1292.	1091	100
995	gi13879442	Mus musculus	Similar to RIKEN cDNA 2310035M22 gene	1056	93
996	gi6692607	Mus musculus	MGA protein	3446	77
996	gi3043560	Homo sapiens	mRNA for KIAA0518 protein, partial cds.	3272	100
996	AAB94560	Homo sapiens	HELI- Human protein sequence SEQ ID NO:15332.	1933	99
997	gi5771392	Homo sapiens	RAB-like protein 2B (RABL2B) mRNA, complete cds.	718	99
997	gi15928838	Homo sapiens	RAB, member of RAS oncogene family-like 2B, clone MGC:10160 IMAGE:3906749, mRNA, complete cds.	718	99
997	gi7263961	Homo sapiens	Human DNA sequence from clone RP11-395L14. Contains (part of) up to six novel genes or pseudogenes, the gene for a novel forkhead protein similar to FOXD4 (forkhead box D4, FREAC5), the gene for a novel phosphoglucomutase like protein, a pseudogene similar to part of DEAD/H (Asp-Glu-Ala-Asp/His) box (S.cerevisiae CHL1-like helicase), an RPL23A (60S ribosomal protein L23A) pseudogene, the RABL2A gene for RAB-like 2A, the gene for a novel protein similar to small nuclear ribonucleoprotein polypeptide A' (SNRPA1) and the 3' part of the gene for a novel protein similar to acrosin (ACR). Contains ESTs, STSs, GSSs and nine putative CpG islands, complete sequence.	714	97
998	gi10440202	Homo sapiens	cDNA: FLJ23495 fis, clone LNG02228.	2398	99
998	AAU17289	Homo sapiens	HUMA- Novel signal transduction pathway protein, Seq ID 854.	487	97
998	AAM92681	Homo sapiens	HUMA- Human digestive system antigen SEQ ID NO: 2030.	487	97
999	gi12653249	Homo sapiens	Similar to CAAX box 1, clone MGC:8471 IMAGE:2821721, mRNA, complete cds.	450	100

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
999	AAY32207	Homo sapiens	INCY- Human receptor molecule (REC) encoded by Incyte clone 2936050.	429	95
999	gi757623	Homo sapiens	Human DNA sequence from clone RP4-809E13 on chromosome Xq26.1-27.1. Contains the gene for a putative prenylated protein, two putative prenylated protein pseudogenes, ESTs, STSs, GSSs and three putative CpG islands, complete sequence.	397	87
1000	gi15778556	Homo sapiens	alpha-1-B glycoprotein precursor (A1BG) mRNA, complete cds.	1487	98
1000	gi11877348	Rattus norvegicus	putative alpha 1B-glycoprotein	518	40
1000	AAY64670	Homo sapiens	GEST Human 5' EST related polypeptide SEQ ID NO:831.	430	76
1001	AAY87315	Homo sapiens	INCY- Human signal peptide containing protein HSPP-92 SEQ ID NO:92.	2817	100
1001	AAM93793	Homo sapiens	HELI- Human polypeptide, SEQ ID NO: 3821.	2527	99
1001	gi12804527	Homo sapiens	hypothetical protein FLJ22405, clone MGC:2543 IMAGE:2961594, mRNA, complete cds.	2194	100
1002	gi4160304	Mus musculus	HS1 binding protein 3	1449	75
1002	AAB95801	Homo sapiens	HELI- Human protein sequence SEQ ID NO:18781.	1082	100
1002	gi10436660	Homo sapiens	cDNA FLJ14249 fis, clone OVARC1001200, weakly similar to Mus musculus mRNA for HS1 binding protein 3.	1082	100
1003	AAY87315	Homo sapiens	INCY- Human signal peptide containing protein HSPP-92 SEQ ID NO:92.	1837	100
1003	gi12804527	Homo sapiens	hypothetical protein FLJ22405, clone MGC:2543 IMAGE:2961594, mRNA, complete cds.	1837	100
1003	gi10438780	Homo sapiens	cDNA: FLJ22405 fis, clone HRC08294.	1837	100
1004	AAM93793	Homo sapiens	HELI- Human polypeptide, SEQ ID NO: 3821.	3401	99
1004	gi17390694	Mus musculus	Similar to hypothetical protein FLJ22405	2543	90
1004	AAY87315	Homo sapiens	INCY- Human signal peptide containing protein HSPP-92 SEQ ID NO:92.	2535	100
1005	AAY87327	Homo sapiens	INCY- Human signal peptide containing protein HSPP-104 SEQ ID NO:104.	584	100
1005	AAY59705	Homo sapiens	GEST Secreted protein 51-41-1-F10-FL1.	554	95

SEQ ID	Hit ID	Species	Description	S score	Percent identity
1005	AAY12865	Homo sapiens	GEST Human 5' EST secreted protein SEQ ID NO:455.	208	100
1006	AAY36237	Homo sapiens	HUMA- Human secreted protein encoded by gene 14.	177	100
1007	AAY87310	Homo sapiens	INCY- Human signal peptide containing protein HSPP-87 SEQ ID NO:87.	370	100
1007	AAG77344	Homo sapiens	HUMA- Human colon cancer antigen protein SEQ ID NO:8110.	201	79
1007	gi14198125	Homo sapiens	clone MGC:18053 IMAGE:4148889, mRNA, complete cds.	68	61
1008	AAB94108	Homo sapiens	HELI- Human protein sequence SEQ ID NO:14340.	1844	96
1008	AAU04557	Homo sapiens	GETH Human Stra6 homologue, PRO10282.	1844	96
1008	gi13560966	Homo sapiens	STRA6 isoform 1 mRNA, complete cds, alternatively spliced.	1844	96
1009	AAY38394	Homo sapiens	HUMA- Human secreted protein encoded by gene No. 9.	213	100
1010	gi10432382	Homo sapiens	Human DNA sequence from clone RP4-717I23 on chromosome 1p21.2-22.3 Contains ESTs, STSs and GSSs. Contains part of a novel gene for a protein similar to Xenopus laevis Sojo protein, a novel gene and a 60S ribosomal protein L39 (RPL39) pseudogene, complete sequence.	3267	100
1010	gi5690435	Xenopus laevis	nuclear protein Sojo	1386	44
1010	AAG75036	Homo sapiens	HUMA- Human colon cancer antigen protein SEQ ID NO:5800.	557	98
1011	AAG00517	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 4598.	160	48
1011	AAO02474	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 16366.	153	45
1011	gi854065	Human herpesvirus 6	U88	145	50
1012	AAY19561	Homo sapiens	HUMA- Amino acid sequence of a human secreted protein.	514	100
1012	AAB38157	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 39 SEQ ID NO:96.	70	30
1012	AAU04958	Homo sapiens	GETH Human Interleukin 17 receptor, IL-17RH4.	69	60
1013	AAR15222	Homo sapiens	TEXA Chronic myelogenous leukaemia-derived myeloid-related protein.	635	100
1013	gi32402	Homo sapiens	Human mRNA for HP-1, a member of the corticostatin/defensin family.	493	100
1013	gi292363	Homo sapiens	Human neutrophil peptide-1	493	100

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
			gene, complete cds.		
1014	gi6330176	Homo sapiens	mRNA for KIAA1167 protein, partial cds.	4079	99
1014	gi8980841	Rattus norvegicus	GRIP-associated protein 1 long form	3814	92
1014	gi17389263	Mus musculus	Similar to GRIP-associated protein 1	3646	89
1015	gi10443047	Homo sapiens	Human DNA sequence from clone RP11-465L10 on chromosome 20. Contains 10 CpG islands, ESTs, STSs and GSSs. Contains the gene for a novel protein similar to Drosophila CG11399, the gene for a novel C2H2 type zinc finger protein similar to chicken FZF-1, a Ferritin light polypeptide (FTL) pseudogene, the MMP9 gene for matrix metalloproteinase 9 (gelatinase B, 92kD gelatinase, 92kD type IV collagenase) (CLG4B), a novel gene, the SLC12A5 gene for solute carrier family 12, (potassium-chloride transporter) member 5 (KIAA1176) and the 3' end of gene KIAA1637, complete sequence.	6471	99
1015	gi10438918	Homo sapiens	cDNA: FLJ22504 fis, clone HRC11430.	4392	98
1015	gi984814	Gallus gallus	zinc finger protein	2127	58
1016	AAE06077	Homo sapiens	HUMA- Human gene 37 encoded secreted protein HDPCJ91, SEQ ID NO:139.	267	100
1016	AAY87100	Homo sapiens	HUMA- Human secreted protein sequence SEQ ID NO:139.	267	100
1016	gi12718812	Yarrowia lipolytica	ND3 protein	69	48
1017	AAY86463	Homo sapiens	HUMA- Human gene 47-encoded protein fragment, SEQ ID NO:378.	361	100
1017	AAY86320	Homo sapiens	HUMA- Human secreted protein HPRBC80, SEQ ID NO:235.	361	100
1017	gi7549633	Arabidopsis thaliana	hypothetical protein	70	31
1018	AAM25384	Homo sapiens	HYSE- Human protein sequence SEQ ID NO:899.	1126	100
1018	AAY27581	Homo sapiens	HUMA- Human secreted protein encoded by gene No. 15.	774	100
1018	gi13785618	Mus musculus	sideroflexin 4	660	60
1019	gi452890	Cricetulus migratorius	serum amyloid P; SAP; female protein; FP	158	71
1019	gi387051	Cricetulus longicaudatus	FP	157	71
1019	gi347257	Mesocricetus auratus	serum amyloid P component	157	71

SEQ ID	Hit ID	Species	Description	S score	Percent identity
1020	gi174287 83	Ralstonia solanacearum	PROBABLE NADP-DEPENDENT OXIDOREDUCTASE OXIDOREDUCTASE PROTEIN	68	29
1020	gi151592 26	Agrobacterium tumefaciens str. C58 (Cereon)	AGR_L_1604p	67	28
1020	gi177425 00	Agrobacterium tumefaciens str. C58 (Dupont)	succinoglycan biosynthesis protein	67	28
1021	gi165539 33	Homo sapiens	cDNA FLJ25217 fis, clone REC08938, highly similar to Oryctolagus cuniculus Na+/glucose cotransporter-related protein mRNA.	1477	100
1021	AAE0661 4	Homo sapiens	SAGA Human protein having hydrophobic domain, HP03974.	1394	100
1021	gi152098 08	Homo sapiens	unnamed protein product	1394	100
1022	AAY167 81	Homo sapiens	GEMY Human secreted protein (clone bh157_7).	1258	100
1022	gi126540 11	Homo sapiens	similar to rat nuclear ubiquitous casein kinase 2, clone MGC:5494 IMAGE:3452665, mRNA, complete cds.	1258	100
1022	gi120536 24	Homo sapiens	mRNA for NUCKS protein.	1258	100
1023	AAB832 46	Homo sapiens	MILL- Human FATP1 SEQ ID NO: 47.	3372	100
1023	AAB832 39	Homo sapiens	MILL- Human FATP1 SEQ ID NO: 38.	3372	100
1023	AAB832 34	Homo sapiens	MILL- Human FATP1 SEQ ID NO: 32.	3372	100
1024	gi159299 04	Homo sapiens	Similar to dolichyl-phosphate mannosyltransferase polypeptide 2, regulatory subunit, clone MGC:21559 IMAGE:4181887, mRNA, complete cds.	366	100
1024	AAY579 05	Homo sapiens	INCY- Human transmembrane protein HTMPN-29.	153	100
1024	gi379036 3	Homo sapiens	mRNA for DPM2, complete cds.	153	100
1025	AAY257 32	Homo sapiens	HUMA- Human secreted protein encoded from gene 22.	212	100
1026	AAG770 16	Homo sapiens	HUMA- Human colon cancer antigen protein SEQ ID NO:7780.	291	100
1026	AAS0319 3_aa1	Homo sapiens	GEHO Human lymphocyte cell surface antigen CD53 cDNA sequence.	116	95
1026	AAV812 20_aa1	Homo sapiens	GEHO Human CD53 antigen cDNA.	116	95
1027	AAY118 48	Homo sapiens	GEST Human 5' EST secreted protein SEQ ID No: 448.	193	100
1027	AAY359 56	Homo sapiens	GEST Extended human secreted protein sequence, SEQ ID NO.	193	100

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
			205.		
1027	AAY360 98	Homo sapiens	GEST Extended human secreted protein sequence, SEQ ID NO. 483.	193	100
1028	AAM939 42	Homo sapiens	HELI- Human polypeptide, SEQ ID NO: 4126.	1354	99
1028	ABB1142 2	Homo sapiens	HYSE- Human Zn finger protein homologue, SEQ ID NO:1792.	953	92
1028	gi646720 6	Homo sapiens	GIOT-4 mRNA for gonadotropin inducible transcription repressor-4, complete cds.	883	56
1029	gi128352 22	Mus musculus	putative	1562	82
1029	gi263665 4	Homo sapiens	Human myosin binding protein H (MyBP-H) gene, complete cds.	1206	66
1029	gi154681 3	Mus musculus	myosin binding protein H	1203	66
1030	AAY363 32	Homo sapiens	HUMA- Human secreted protein encoded by gene 109.	268	100
1030	gi100389 17	Buchnera sp. APS	hypothetical protein	76	42
1030	gi750051 8	Unknown	hypothetical protein F35E2.7 - Caenorhabditis elegans >	63	38
1031	gi971940 9	Homo sapiens	candidate tumor suppressor protein mRNA, complete cds.	2030	99
1031	gi128565 14	Mus musculus	putative	922	86
1031	AAU220 41	Homo sapiens	HUMA- Human cardiovascular system antigen polypeptide SEQ ID No 815.	703	92
1032	AAR152 22	Homo sapiens	TEXA Chronic myelogenous leukaemia-derived myeloid-related protein.	635	100
1032	gi32402	Homo sapiens	Human mRNA for HP-1, a member of the corticostatin/defensin family.	493	100
1032	gi292363	Homo sapiens	Human neutrophil peptide-1 gene, complete cds.	493	100
1033	gi165525 02	Homo sapiens	cDNA FLJ32395 fis, clone SKMUS2000117, moderately similar to Homo sapiens MAGEF1 mRNA.	1599	100
1033	gi126591 42	Mus musculus	mage-g1	1178	76
1033	gi128571 18	Mus musculus	putative	1178	76
1034	AAB496 50	Homo sapiens	CURA- Human SEC2 protein sequence SEQ ID 4.	2615	100
1034	gi122265 32	Homo sapiens	unnamed protein product	2615	100
1034	gi147148 86	Mus musculus	Unknown (protein for IMAGE:3498778)	2343	89
1035	AAM237 21	Homo sapiens	HYSE- Human EST encoded protein SEQ ID NO: 1246.	2889	100
1035	AAM792	Homo sapiens	HYSE- Human protein SEQ ID	2676	100

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
	29		NO 1891.		
1035	AAB883 70	Homo sapiens	HELI- Human membrane or secretory protein clone PSEC0106.	1611	100
1036	AAY596 57	Homo sapiens	GEST Secreted protein 108-003-5-0-A8-FL.	689	100
1036	gi144956 99	Homo sapiens	clone MGC:15961 IMAGE:3538818, mRNA, complete cds.	689	100
1036	gi144245 22	Homo sapiens	clone MGC:14327 IMAGE:4298098, mRNA, complete cds.	689	100
1037	AAY276 26	Homo sapiens	HUMA- Human secreted protein encoded by gene No. 60.	352	100
1038	AAD183 56_aa1	Homo sapiens	INCY- Human lipid metabolism enzyme-5 (LME-5) cDNA.	1748	100
1038	AAB735 60	Homo sapiens	MILL- Human lipase 18892.	1748	100
1038	AAE1099 6	Homo sapiens	INCY- Human lipid metabolism enzyme-5 (LME-5) protein.	1748	100
1039	AAG034 75	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 7556.	448	100
1039	AAY128 61	Homo sapiens	GEST Human 5' EST secreted protein SEQ ID NO:451.	448	100
1039	gi56760	Rattus norvegicus	neuronal nonacetlycholine binding subunit	75	23
1040	AAY530 49	Homo sapiens	GEMY Human secreted protein clone cj378_3 protein sequence SEQ ID NO:104.	463	100
1040	gi136036 74	Stellilabium pogonostalix	maturase	78	38
1040	gi136036 76	Telipogon parvulus	maturase	74	36
1041	AAY413 54	Homo sapiens	HUMA- Human secreted protein encoded by gene 47 clone HUFCJ30.	288	100
1041	gi152304 14	Arabidopsis thaliana	putative protein	63	43
1042	AAW747 77	Homo sapiens	HUMA- Human secreted protein encoded by gene 48 clone HFCAI74.	245	100
1042	gi154889 20	Homo sapiens	Similar to RIKEN cDNA 2010107G23 gene, clone MGC:9596 IMAGE:3896656, mRNA, complete cds.	245	100
1042	gi128424 65	Mus musculus	putative	241	97
1043	gi152780 28	Homo sapiens	beta-galactose-3-O-sulfotransferase, 4, clone MGC:15045 IMAGE:3636329, mRNA, complete cds.	2568	100
1043	gi147945 24	Homo sapiens	Galbeta1-3GalNAc 3'-sulfotransferase mRNA, complete cds.	2564	99
1043	AAB938 92	Homo sapiens	HELI- Human protein sequence SEQ ID NO:13832.	2556	99
1044	gi153215	Homo sapiens	empty spiracles-like protein	1341	100

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
	90		(EMX2) mRNA, complete cds.		
1044	gi132767 73	Homo sapiens	mRNA; cDNA DKFZp761M1614 (from clone DKFZp761M1614).	1341	100
1044	gi165496 86	Homo sapiens	cDNA FLJ30479 fis, clone BRAWH1000168, highly similar to Homeotic protein emx2.	1336	99
1045	gi168770 66	Homo sapiens	clone MGC:24447 IMAGE:4077762, mRNA, complete cds.	362	100
1045	gi168770 59	Homo sapiens	clone MGC:24437 IMAGE:4075637, mRNA, complete cds.	362	100
1045	AAY949 59	Homo sapiens	GEMY Human secreted protein clone mc300_1 protein sequence SEQ ID NO:124.	204	97
1046	gi239445 1	Caenorhabditis elegans	Hypothetical protein ZC178.2	406	30
1046	AAB875 75	Homo sapiens	GETH Human PRO1342.	384	38
1046	AAY994 08	Homo sapiens	GETH Human PRO1342 (UNQ697) amino acid sequence SEQ ID NO:243.	384	38
1047	gi120531 47	Homo sapiens	mRNA; cDNA DKFZp434F1726 (from clone DKFZp434F1726).	1484	98
1047	ABB1173 9	Homo sapiens	HYSE- Human IF-gamma receptor homologue, SEQ ID NO:2109.	1044	100
1047	AAR049 32	Homo sapiens	YEDA Interferon-gamma receptor segment from clone 39 responsiblefor binding the target.	829	98
1048	AAG739 89	Homo sapiens	HUMA- Human colon cancer antigen protein SEQ ID NO:4753.	957	100
1048	AAB589 98	Homo sapiens	HUMA- Breast and ovarian cancer associated antigen protein sequence SEQ ID 706.	957	100
1048	AAM891 00	Homo sapiens	HUMA- Human immune/haematopoietic antigen SEQ ID NO:16693.	362	67
1049	gi100473 33	Homo sapiens	mRNA for KIAA1628 protein, partial cds.	5204	100
1049	gi118629 39	Mus musculus	DDM36	4378	89
1049	gi118629 41	Mus musculus	DDM36E	4366	88
1050	gi190647	Homo sapiens	Human pregnancy-specific beta-1 glycoprotein (PSG) mRNA, complete cds.	611	72
1050	gi984306	Homo sapiens	Human pregnancy-specific glycoprotein 13 (PSG13') mRNA, complete cds.	606	71
1050	gi190568	Homo sapiens	Human pregnancy-specific beta-1-glycoprotein 11 (PSG11) mRNA, complete cds.	585	64
1051	AAM436	Homo sapiens	HUMA- Human polypeptide	588	100

SEQ ID	Hit ID	Species	Description	S score	Percent identity
	55		SEQ ID NO 333.		
1051	AAM43588	Homo sapiens	HUMA- Human polypeptide SEQ ID NO 266.	588	100
1051	AAW60043	Homo sapiens	HUTC- Human MHC class I chain-related gene A (MICA) polypeptide.	588	100
1052	gi4877836	Rattus norvegicus	TRP2	524	74
1052	gi11095641	Mus musculus	transient receptor potential channel 2-beta	521	73
1052	gi11095639	Mus musculus	transient receptor potential channel 2-alpha	521	73
1053	AAB94188	Homo sapiens	HELI- Human protein sequence SEQ ID NO:14511.	1209	97
1053	gi10434224	Homo sapiens	cDNA FLJ12623 fis, clone NT2RM4001746.	1209	97
1053	gi12652797	Homo sapiens	clone MGC:5179 IMAGE:2900118, mRNA, complete cds.	1066	88
1054	AAY38389	Homo sapiens	HUMA- Human secreted protein encoded by gene No. 4.	152	90
1054	gi13624635	Euglena viridis	maturase-like protein	63	42
1055	AAY27582	Homo sapiens	HUMA- Human secreted protein encoded by gene No. 16.	320	100
1055	gi13421631	Caulobacter crescentus	conserved hypothetical protein	65	27
1055	gi16124713	Caulobacter crescentus] > [Caulobacter crescentus	conserved hypothetical protein	65	27
1056	AAO08759	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 22651.	613	100
1056	gi10437569	Homo sapiens	cDNA: FLJ21463 fis, clone COL04765.	269	57
1056	AAY45382	Homo sapiens	HUMA- Human secreted protein fragment encoded from gene 28.	266	58
1057	AAE05175	Homo sapiens	INCY- Human drug metabolising enzyme (DME-6) protein.	1830	99
1057	AAU12225	Homo sapiens	GETH Human PRO4404 polypeptide sequence.	1830	99
1057	AAU18363	Homo sapiens	HUMA- Human endocrine polypeptide SEQ ID No 318.	1092	95
1058	AAG81274	Homo sapiens	ZYMO Human AFP protein sequence SEQ ID NO:66.	815	96
1058	gi14035856	Homo sapiens	unnamed protein product	815	96
1058	AAG81273	Homo sapiens	ZYMO Human AFP protein sequence SEQ ID NO:64.	652	99
1059	AAY35980	Homo sapiens	GEST Extended human secreted protein sequence, SEQ ID NO. 229.	581	97
1059	ABB11960	Homo sapiens	HYSE- Human neuroendocrine-specific protein-like homologue, SEQ ID NO:2330.	246	100
1059	AAZ3831	Homo sapiens	PROT- Human transmembrane	240	97

SEQ ID	Hit ID	Species	Description	S score	Percent identity
	9_aa1		protein cDNA clone HP02061.		
1060	gi795932 5	Homo sapiens	mRNA for KIAA1529 protein, partial cds.	8481	100
1060	gi128363 54	Mus musculus	putative	511	63
1060	AAW036 26	Homo sapiens	UYNY Human thyrotropin GPR N-terminal sequence.	236	31
1061	AAY762 00	Homo sapiens	HUMA- Human secreted protein encoded by gene 77.	262	100
1061	gi159243 50	Staphylococcus aureus subsp. aureus Mu50	oxacillin resistance-related FmtC protein	64	31
1061	gi124836 31	Staphylococcus aureus	FmtC	64	31
1062	AAY362 70	Homo sapiens	HUMA- Human secreted protein encoded by gene 47.	359	100
1062	gi499607 9	Human herpesvirus 6	64% identical to U95 gene of strain U1102 of HHV-6~MCMV IE2 homolog, US22 gene family	68	37
1062	gi573357 6	Human herpesvirus 6B	U95	66	37
1063	gi526274 8	Rattus norvegicus	Proline rich synapse associated protein 2	3930	93
1063	gi738105 6	Rattus norvegicus	Shank postsynaptic density protein 3a	3895	92
1063	gi133591 73	Homo sapiens	mRNA for KIAA1650 protein, partial cds.	3085	100
1064	gi143367 49	Homo sapiens	16p13.3 sequence section 6 of 8.	974	99
1064	gi104400 21	Homo sapiens	cDNA: FLJ23360 fis, clone HEP15172.	974	99
1064	AAB941 88	Homo sapiens	HELI- Human protein sequence SEQ ID NO:14511.	914	78
1065	AAB941 88	Homo sapiens	HELI- Human protein sequence SEQ ID NO:14511.	1227	100
1065	gi104342 24	Homo sapiens	cDNA FLJ12623 fis, clone NT2RM4001746.	1227	100
1065	gi126527 97	Homo sapiens	clone MGC:5179 IMAGE:2900118, mRNA, complete cds.	1084	90
1066	AAY824 88	Homo sapiens	NISC- Human L-type amino acid transporter 1 protein sequence SEQ ID NO:2.	2438	94
1066	gi592673 2	Homo sapiens	mRNA for L-type amino acid transporter 1, complete cds.	2438	94
1066	gi442664 0	Homo sapiens	L-type amino acid transporter subunit LAT1 mRNA, complete cds.	2438	94
1067	AAG813 26	Homo sapiens	ZYMO Human AFP protein sequence SEQ ID NO:170.	1135	100
1067	gi140359 60	Homo sapiens	unnamed protein product	1135	100
1067	AAY788 05	Homo sapiens	PROT- Hydrophobic domain containing protein clone HP10508 protein sequence.	1053	99
1068	gi120529 83	Homo sapiens	mRNA; cDNA DKFZp434I1610 (from clone DKFZp434I1610);	2502	63

SEQ ID	Hit ID	Species	Description	S score	Percent identity
			complete cds.		
1068	AAM79760	Homo sapiens	HYSE- Human protein SEQ ID NO 3406.	2270	61
1068	AAB94388	Homo sapiens	HELI- Human protein sequence SEQ ID NO:14947.	2050	57
1069	gi12655091	Homo sapiens	AD-003 protein, clone MGC:783 IMAGE:3050940, mRNA, complete cds.	546	53
1069	gi6523799	Homo sapiens	adrenal gland protein AD-003 mRNA, complete cds.	536	53
1069	AAM51824	Homo sapiens	BIOW- Human transcription regulator 13.	370	61
1070	gi14328009	Homo sapiens	clone IMAGE:3942111, mRNA, partial cds.	2392	100
1070	gi14585869	Homo sapiens	hypothetical protein SB146	2389	99
1070	gi16041767	Homo sapiens	Similar to NADPH oxidase-related, C2 domain-containing protein, clone MGC:23187 IMAGE:4851468, mRNA, complete cds.	2384	99
1071	ABB12245	Homo sapiens	HYSE- Human CKSR-2 homologue, SEQ ID NO:2615.	256	50
1071	AAZ88239_aa1	Homo sapiens	INCY- Human cytokine signal regulator CKSR-2 encoding cDNA SEQ ID NO:4.	245	51
1071	AAB67667	Homo sapiens	INCY- Amino acid sequence of human cytokine signal regulator 2.	245	51
1072	gi15418997	Homo sapiens	capillary morphogenesis protein-1 mRNA, complete cds.	3015	100
1072	AAB95505	Homo sapiens	HELI- Human protein sequence SEQ ID NO:18062.	2163	99
1072	gi10435718	Homo sapiens	cDNA FLJ13645 fis, clone PLACE1011310, weakly similar to MYOSIN HEAVY CHAIN, GIZZARD SMOOTH MUSCLE.	2163	99
1073	gi15985082	Homo sapiens	unnamed protein product	3158	99
1073	AAU04958	Homo sapiens	GETH Human Interleukin 17 receptor, IL-17RH4.	3148	97
1073	gi6453552	Homo sapiens	mRNA; cDNA DKFZp434N1928 (from clone DKFZp434N1928).	3007	100
1074	AAB93827	Homo sapiens	HELI- Human protein sequence SEQ ID NO:13641.	884	99
1074	AAB23606	Homo sapiens	ALPH- Human secreted protein SEQ ID NO: 12.	884	99
1074	gi10433126	Homo sapiens	cDNA FLJ11790 fis, clone HEMBA1006091.	884	99
1075	gi15777931	Homo sapiens	DGCRK3 gene for G-protein beta subunit like protein, complete cds.	1731	99
1075	gi15082309	Homo sapiens	clone MGC:19898 IMAGE:4548339, mRNA, complete cds.	1731	99

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
1075	gi13359165	Homo sapiens	mRNA for KIAA1645 protein, partial cds.	1731	99
1076	gi16605905	Homo sapiens	unnamed protein product	100	32
1076	gi11611188	Homo sapiens	Human DNA sequence from clone RP4-688G8 on chromosome 20q11.2-12. Contains the gene for a novel protein similar to ribosomal protein S2 (RPS2), a gene encoding a protein similar to basic protease inhibitor chelonianin, a novel gene, the 3' end of a novel gene, ESTs, STSs, GSSs and a CpG island, complete sequence.	94	32
1076	AAY35935	Homo sapiens	GEST Extended human secreted protein sequence, SEQ ID NO. 184.	83	29
1077	gi16554229	Homo sapiens	cDNA FLJ25436 fis, clone TST08261.	1164	100
1077	gi15082426	Homo sapiens	Similar to RIKEN cDNA 2810055F11 gene, clone MGC:20203 IMAGE:4684687, mRNA, complete cds.	1156	99
1077	gi12858155	Mus musculus	putative	1054	89
1078	gi15559290	Homo sapiens	clone MGC:20275 IMAGE:3842589, mRNA, complete cds.	1917	100
1078	gi15625564	Homo sapiens	WD40- and FYVE-domain containing protein 2 (WDF2) mRNA, complete cds.	1893	99
1078	gi16554204	Homo sapiens	cDNA FLJ25420 fis, clone TST03665.	1380	100
1079	AAB92775	Homo sapiens	HELI- Human protein sequence SEQ ID NO:11256.	3212	99
1079	gi14042169	Homo sapiens	cDNA FLJ14564 fis, clone NT2RM4000229, weakly similar to Gallus gallus actin filament-associated protein (AFAP-110) mRNA.	3212	99
1079	gi15620887	Homo sapiens	mRNA for KIAA1914 protein, partial cds.	1702	100
1080	AAG89172	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 292.	591	100
1080	AAY12532	Homo sapiens	GEST Human 5' EST secreted protein SEQ ID NO: 197 from WO 9906553.	512	91
1080	AAB87173	Homo sapiens	MILL- Human secreted protein TANGO 402 S22T variant, SEQ ID NO:215.	119	44
1081	gi9963804	Homo sapiens	zinc finger protein ZNF286 (ZNF286) mRNA, complete cds.	574	94
1081	gi14017965	Homo sapiens	mRNA for KIAA1874 protein, partial cds.	517	91
1081	AAU16238	Homo sapiens	HUMA- Human novel secreted protein, Seq ID 1191.	362	100

SEQ ID	Hit ID	Species	Description	S score	Percent identity
1082	AAG03810	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 7891.	841	99
1082	gi186800	Homo sapiens	Human ribosomal protein L12 mRNA, complete cds.	841	99
1082	gi14198333	Homo sapiens	ribosomal protein L12, clone MGC:9760 IMAGE:3855674, mRNA, complete cds.	841	99
1083	AAY12902	Homo sapiens	GEST Human 5' EST secreted protein SEQ ID NO:492.	134	100
1084	gi186774	Homo sapiens	Human Kruppel related zinc finger protein (HTF10) mRNA, complete cds.	3082	57
1084	gi2739353	Homo sapiens	DNA from chromosome 19, BAC 33152, complete sequence.	2985	55
1084	AAM79739	Homo sapiens	HYSE- Human protein SEQ ID NO 3385.	2612	53
1085	AAV48125_aa1	Homo sapiens	HUMA- Nucleotide sequence encoding clone HMWGS46 of Prohibitin receptor family.	1354	93
1085	ABB11913	Homo sapiens	HYSE- Human B-cell receptor associated protein homologue, SEQ ID NO:2283.	1354	93
1085	AAY94443	Homo sapiens	UNII Human repressor of estrogen repressor activity (REA) protein.	1354	93
1086	AAG72370	Homo sapiens	YEDA Human OR-like polypeptide query sequence, SEQ ID NO: 2051.	333	100
1086	AAG71453	Homo sapiens	YEDA Human olfactory receptor polypeptide, SEQ ID NO: 1134.	333	100
1086	AAE04556	Homo sapiens	INCY- Human G-protein coupled receptor-12 (GCREC-12) protein.	315	100
1087	AAG81323	Homo sapiens	ZYMO Human AFP protein sequence SEQ ID NO:164.	792	100
1087	gi14035954	Homo sapiens	unnamed protein product	792	100
1087	gi12838799	Mus musculus	putative	564	76
1088	AAE09691	Homo sapiens	HUMA- Human gene 2 encoding novel protein HCOKA10, SEQ ID NO:38.	96	55
1088	AAG76125	Homo sapiens	HUMA- Human colon cancer antigen protein SEQ ID NO:6889.	96	55
1088	AAU16944	Homo sapiens	HUMA- Human novel secreted protein, SEQ ID 185.	96	55
1089	AAY13037	Homo sapiens	GEST Human secreted protein encoded by 5' EST SEQ ID NO: 51.	187	100
1089	AAY36395	Homo sapiens	HUMA- Fragment of human secreted protein encoded by gene 5.	79	39
1089	gi13096804	Mus musculus	Unknown (protein for IMAGE:3586067)	77	40
1090	gi151478	Mus musculus	Spred-2	2098	92

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
	77				
1090	gi15147875	Mus musculus	Spred-1	1101	52
1090	AAU17415	Homo sapiens	HUMA- Novel signal transduction pathway protein, Seq ID 980.	1029	98
1091	gi16551762	Homo sapiens	cDNA FLJ31812 fis, clone NT2RI2009406, moderately similar to Homo sapiens rec mRNA.	804	97
1091	gi6682873	Homo sapiens	rec mRNA, complete cds.	453	55
1091	gi7230612	Rattus norvegicus	small rec	451	56
1092	gi12224885	Homo sapiens	mRNA; cDNA DKFZp761I1011 (from clone DKFZp761I1011).	241	94
1092	gi3880560	Caenorhabditis elegans	Similarity to Yeast E1-E2 ATPase (SW:YED1_YEAST), contains similarity to Pfam domain: PF00122 (E1-E2 ATPase), Score=102.4, E-value=2.7e-28, N=4~cDNA EST yk5f9.5 comes from this gene~cDNA EST yk10d12.5 comes from this gene~cDNA EST yk5f9.3 comes from this gene~cDNA EST yk10d12.3 comes from this gene~cDNA EST yk40h11.5 comes from this gene~cDNA EST yk131g11.3 comes from this gene~cDNA EST yk131g11.5 comes from this gene~cDNA EST yk133d6.5 comes from this gene~cDNA EST yk318f2.3 comes from this gene~cDNA EST yk318f2.5 comes from this gene~cDNA EST yk122a12.5 comes from this gene~cDNA EST yk248h3.5 comes from this gene~cDNA EST yk260b8.5 comes from this gene~cDNA EST yk286h5.5 comes from this gene~cDNA EST yk356g1.5 comes from this gene	135	50
1092	gi16516658	Homo sapiens	ORF for hypothetical protein.	129	52
1093	AAB40996	Homo sapiens	CURA- Human ORFX ORF760 polypeptide sequence SEQ ID NO:1520.	349	40
1093	gi16516658	Homo sapiens	ORF for hypothetical protein.	349	40
1093	gi10436963	Homo sapiens	cDNA: FLJ20986 fis, clone CAE01156.	349	40
1094	gi16904208	Homo sapiens	very large G protein-coupled receptor 1b (VLGR1) mRNA, complete cds.	1014	96
1094	gi169042	Mus musculus	very large G protein-coupled	808	78

SEQ ID	Hit ID	Species	Description	S score	Percent identity
	10		receptor 1		
1094	gi120444 71	Homo sapiens	mRNA; cDNA DKFZp761P0710 (from clone DKFZp761P0710); complete cds.	173	27
1095	gi124839 02	Rattus norvegicus	zinc finger protein HIT-10	1545	47
1095	AAB958 62	Homo sapiens	HELI- Human protein sequence SEQ ID NO:18929.	1234	50
1095	gi104367 89	Homo sapiens	cDNA FLJ14345 fis, clone THYRO1001189, weakly similar to ZINC FINGER PROTEIN 91.	1234	50
1096	AAB509 63	Homo sapiens	GETH Human PRO1286 protein.	466	100
1096	AAU124 21	Homo sapiens	GETH Human PRO1286 polypeptide sequence.	466	100
1096	AAU091 79	Homo sapiens	GETH Human PRO1268 polypeptide.	466	100
1097	AAE1202 3	Homo sapiens	INCY- Human G-protein coupled receptor, GCREC-2.	2849	98
1097	AAG681 26	Homo sapiens	FARB Human 7TM-GPCR protein sequence SEQ ID NO:6.	2824	98
1097	gi175125 39	Mus musculus	Unknown (protein for MGC:29266)	2183	73
1098	AAB941 08	Homo sapiens	HELI- Human protein sequence SEQ ID NO:14340.	273	100
1098	AAU045 58	Homo sapiens	GETH Human Stra6 homologue, PRO19578.	273	100
1098	AAU045 57	Homo sapiens	GETH Human Stra6 homologue, PRO10282.	273	100
1099	AAU123 82	Homo sapiens	GETH Human PRO792 polypeptide sequence.	137	32
1099	AAB244 16	Homo sapiens	GETH Human PRO792 protein sequence SEQ ID NO:155.	137	32
1099	AAB240 55	Homo sapiens	GETH Human PRO792 protein sequence SEQ ID NO:31.	137	32
1100	gi633042 2	Homo sapiens	mRNA for KIAA1202 protein, partial cds.	4913	99
1100	gi123140 62	Homo sapiens	Human DNA sequence from clone RP11-119E20 on chromosome Xp11.21-11.23 Contains part of the gene for KIAA1202 protein, ESTs, STSs and GSSs, complete sequence.	4696	99
1100	gi154212 01	Homo sapiens	SHAP-A (SHAP) mRNA, partial cds, alternatively spliced.	3845	99
1101	gi108346 07	Homo sapiens	cadherin 20 (CDH20) mRNA, complete cds.	4170	99
1101	gi410175 1	Mus musculus	cadherin 7 precursor	4032	96
1101	gi854635	Xenopus laevis	F-cadherin	3251	78
1102	AAY363 10	Homo sapiens	HUMA- Human secreted protein encoded by gene 87.	250	100
1103	gi165514 23	Homo sapiens	cDNA FLJ31547 fis, clone NT2RI2001010, weakly similar to FATTY ACYL-COA	853	98

SEQ ID	Hit ID	Species	Description	S score	Percent identity
			HYDROLASE PRECURSOR, MEDIUM CHAIN (EC 3.1.2.14).		
1103	gi2641986	Mesocricetus auratus	carboxylesterase precursor	438	50
1103	AAD00680_aa1	Homo sapiens	INCY- Human Hydrolase protein-5 (HYDRL-5) encoding cDNA.	428	52
1104	AAE04836	Homo sapiens	SUGE- Human SGP018 phosphatase polypeptide.	4915	95
1104	gi12718835	Homo sapiens	unknown mRNA.	3350	99
1104	AAB40919	Homo sapiens	CURA- Human ORFX ORF683 polypeptide sequence SEQ ID NO:1366.	718	79
1105	gi13492650	Rattus norvegicus	potassium channel beta subunit KChIP4	1284	99
1105	gi11527318	Mus musculus	calsenilin-like protein	1281	99
1105	gi14091336	Homo sapiens	KCHIP4.1 (KCNIP4) mRNA, complete cds, alternatively spliced.	1278	99
1106	AAY13126	Homo sapiens	GEST Human secreted protein encoded by 5' EST SEQ ID NO: 140.	160	96
1107	gi4126593	Cyprinus carpio	complement C3-S	156	29
1107	gi4126589	Cyprinus carpio	complement C3-H1	148	26
1107	gi305335	Cavia porcellus	complement C3 protein (GPC3) precursor	146	50
1108	gi11244873	Homo sapiens	PR-domain-containing protein 16 (PRDM16) mRNA, complete cds.	6646	99
1108	gi12697895	Homo sapiens	mRNA for KIAA1675 protein, partial cds.	3570	99
1108	gi545408	human, leukemic cell line SKH1, mRNA Mutant, 5938 nt]. [Homo sapiens	AML1-EVI-1=AML1-EVI-1 fusion protein {rearranged translocation}	3181	53
1109	gi10732815	Homo sapiens	concentrative Na+-nucleoside cotransporter hCNT3 (CNT3) mRNA, complete cds.	3609	100
1109	gi10732817	Mus musculus	concentrative Na+-nucleoside cotransporter mCNT3	2872	78
1109	gi15489379	Mus musculus	solute carrier family 28 (sodium-coupled nucleoside transporter), member 3	2859	77
1110	gi17865150	Plasmodium berghei	cysteine repeat modular protein 3 PbCRM3	97	30
1110	gi6650711	Giardia intestinalis	variant-specific surface protein VSP1267-2	96	29
1110	gi861294	Caenorhabditis elegans	F35D2.4 gene product	94	31
1111	gi7959177	Homo sapiens	mRNA for KIAA1458 protein, partial cds.	2993	100

SEQ ID	Hit ID	Species	Description	S score	Percent identity
1111	AAB94791	Homo sapiens	HELI- Human protein sequence SEQ ID NO:15907.	1037	100
1111	AAU01101	Homo sapiens	HUMA- Gene 38 Human secreted protein homologous amino acid sequence.	842	99
1112	gi17390760	Mus musculus	RIKEN cDNA 2610205H19 gene	664	99
1112	gi12848292	Mus musculus	putative	664	99
1112	gi203072	Rattus sp.	0-44 protein	661	98
1113	gi14250319	Homo sapiens	clone IMAGE:3448367, mRNA, partial cds.	2143	93
1113	gi14582773	Homo sapiens	sumo/sentrin-specific protease	2138	93
1113	gi17026032	Macaca fascicularis	hypothetical protein	2068	89
1114	AAB93777	Homo sapiens	HELI- Human protein sequence SEQ ID NO:13502.	1064	99
1114	AAM41387	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 6318.	1064	99
1114	AAM39601	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 2746.	1064	99
1115	gi15277240	Homo sapiens	genomic DNA, chromosome 6p21.3, HLA Class I region, section 17/20.	2256	100
1115	gi11875405	Homo sapiens	HZFw1 protein mRNA, complete cds.	2251	99
1115	gi11875407	Homo sapiens	HZFw2 protein mRNA, complete cds.	1733	99
1116	AAB95726	Homo sapiens	HELI- Human protein sequence SEQ ID NO:18602.	1644	99
1116	AAB95109	Homo sapiens	HELI- Human protein sequence SEQ ID NO:17089.	1644	99
1116	gi14042078	Homo sapiens	cDNA FLJ14510 fis, clone NT2RM1000623, weakly similar to RIBONUCLEASE INHIBITOR.	1644	99
1117	gi14009346	Homo sapiens	nGAP-like protein (AF9q34) mRNA, complete cds.	5475	98
1117	gi15277525	Rattus norvegicus	DOC2/DAB2 interactive protein	5006	96
1117	gi12698031	Homo sapiens	mRNA for KIAA1743 protein, partial cds.	3024	98
1118	AAB65211	Homo sapiens	GETH Human PRO1152 (UNQ582) protein sequence SEQ ID NO:216.	1937	99
1118	AAB68883	Homo sapiens	INCY- Human RECAP polypeptide, SEQ ID NO: 13.	1937	99
1118	AAU28183	Homo sapiens	HYSE- Novel human secretory protein, Seq ID No 352.	1937	99
1119	AAG02197	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 6278.	522	99
1119	AAU17224	Homo sapiens	HUMA- Novel signal transduction pathway protein, Seq ID 789.	485	100
1119	AAU17597	Homo sapiens	HUMA- Novel signal transduction pathway protein,	449	100

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
			Seq ID 1162.		
1120	gi179636	Homo sapiens	cytoplasmic phosphotyrosyl protein phosphatase (clone type 1) complete cds.	717	88
1120	gi1147812	Homo sapiens	Human red cell-type low molecular weight acid phosphatase (ACP1) gene, exon 6 and 7, complete cds.	717	88
1120	gi5759131	Rattus norvegicus	low molecular weight protein tyrosine phosphatase isoform A	647	76
1121	gi7670466	Mus musculus	unnamed protein product	255	61
1121	gi3413810	Mus musculus	Bassoon	105	25
1121	gi50715	Mus musculus	myosin heavy chain	103	26
1122	gi10440335	Homo sapiens	cDNA: FLJ23594 fis, clone LNG14867.	1513	100
1122	gi14602889	Homo sapiens	clone MGC:13119 IMAGE:4100726, mRNA, complete cds.	702	53
1122	gi14249827	Homo sapiens	clone MGC:10992 IMAGE:3637387, mRNA, complete cds.	702	53
1123	ABB11200	Homo sapiens	HYSE- Human Kupffer cell receptor homologue, SEQ ID NO:1570.	1838	99
1123	gi15489066	Mus musculus	Kupffer cell c-type lectin receptor	1000	45
1123	gi1669360	Mus musculus	Kupffer cell receptor	1000	45
1124	AAY30847	Homo sapiens	HUMA- Human secreted protein encoded from gene 37.	239	100
1124	gi14596027	Arabidopsis thaliana	Unknown protein	70	45
1124	gi9758957	Arabidopsis thaliana	contains similarity to unknown protein~gb|AAF64546.1~gene_id:MRB17.15	70	45
1125	AAY01390	Homo sapiens	HUMA- Secreted protein encoded by gene 8 clone HTXDJ88.	301	100
1125	gi15639942	Treponema pallidum] > [Treponema pallidum	dicarboxylate transporter (dctM)	63	38
1126	AAG68189	Homo sapiens	GENO- Cytosolic thyroid hormone-binding protein SEQ ID NO:105.	152	78
1126	gi35505	Homo sapiens	H.sapiens M gene for M1-type and M2-type pyruvate kinase.	152	78
1126	gi338827	Homo sapiens	Human TCB gene encoding cytosolic thyroid hormone-binding protein, complete cds.	152	78
1127	gi4261689	Homo sapiens	complement factor B mRNA, complete cds.	3976	97
1127	gi297569	Homo sapiens	H.sapiens mRNA for complement factor B.	3976	97
1127	AAX041	Homo sapiens	UNIW Human Factor B	3972	97

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
	25_aa1		encoding cDNA.		
1128	AAB500 12	Homo sapiens	PHAA Wild-type human alpha7 ligand gated ion channel.	1794	100
1128	AAB826 90	Homo sapiens	TEWE- Nicotinic acetylcholine receptor alpha7.	1794	100
1128	AAB240 88	Homo sapiens	GETH Human PRO2145 protein sequence SEQ ID NO:77.	1794	100
1129	gi537329	Homo sapiens	Human (clone pHAIV2-12) alpha-2 collagen type IV (COL4A2) mRNA, 3' end.	3756	99
1129	gi102232 3	Mus musculus	collagen alpha-2(IV) chain	3480	87
1129	gi556299	Mus musculus	alpha-2 type IV collagen	3477	87
1130	gi150114 89	Tetrahymena thermophila	heme maturase	68	29
1131	AAB509 64	Homo sapiens	GETH Human PRO1313 protein.	926	100
1131	AAB472 90	Homo sapiens	GETH PRO1313 polypeptide.	926	100
1131	AAB244 31	Homo sapiens	GETH Human PRO1313 protein sequence SEQ ID NO:216.	926	100
1132	gi128553 07	Mus musculus	putative	2919	89
1132	gi155595 25	Homo sapiens	Similar to RIKEN cDNA 4932416D09 gene, clone IMAGE:4578228, mRNA, partial cds.	2523	99
1132	AAY540 52	Homo sapiens	PHAA An angiogenesis-associated protein which binds plasminogen.	1435	62
1133	AAY130 84	Homo sapiens	GEST Human secreted protein encoded by 5' EST SEQ ID NO: 98.	127	62
1133	gi104404 68	Homo sapiens	mRNA for FLJ00070 protein, partial cds.	75	41
1133	gi455864 0	Homo sapiens	chromosome 19, cosmid R27516, complete sequence.	74	44
1134	AAM802 75	Homo sapiens	HYSE- Human protein SEQ ID NO 3921.	1510	99
1134	AAM792 91	Homo sapiens	HYSE- Human protein SEQ ID NO 1953.	1500	99
1134	gi168774 49	Homo sapiens	hypothetical protein MGC20781, clone MGC:21670 IMAGE:3885455, mRNA, complete cds.	1367	100
1135	gi771086 9	Homo sapiens	Human DNA sequence from clone RP11-31M2 on chromosome 9p23-24.3. Contains (part of) the gene for a novel protein similar to the GLI family of zinc finger proteins, STSs, GSSs and two putative CpG islands, complete sequence.	1629	100
1135	AAM254 73	Homo sapiens	HYSE- Human protein sequence SEQ ID NO:988.	861	98
1135	gi144860 69	Drosophila melanogaster	Zn finger transcription factor lame duck	699	63

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
1136	AAR055 62	Homo sapiens	DANA- Laminin -binding protein encoded by insert from J9 lambda gt10 phage.	1327	94
1136	gi307105	Homo sapiens	Human colin carcinoma laminin-binding protein mRNA, complete cds.	1327	94
1136	gi163076 02	Homo sapiens	laminin receptor 1 (67kD, ribosomal protein SA), clone MGC:17122 IMAGE:3446816, mRNA, complete cds.	1327	94
1137	AAB438 84	Homo sapiens	HUMA- Human cancer associated protein sequence SEQ ID NO:1329.	734	98
1137	gi179281	Homo sapiens	ATP synthase beta subunit precursor (ATPSB) gene, complete cds.	734	98
1137	gi167413 73	Homo sapiens	Similar to ATP synthase, H+ transporting, mitochondrial F1 complex, beta polypeptide, clone MGC:5231 IMAGE:2900336, mRNA, complete cds.	734	98
1138	AAG014 68	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 5549.	282	98
1138	AAG014 67	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 5548.	206	100
1138	AAB438 84	Homo sapiens	HUMA- Human cancer associated protein sequence SEQ ID NO:1329.	206	100
1139	gi179279	Homo sapiens	Human ATP synthase beta subunit gene, exons 8-10.	757	69
1139	gi128456 67	Mus musculus	putative	744	68
1139	gi28940	Homo sapiens	Human mRNA for F1-ATPase beta subunit (F-1 beta).	742	69
1140	AAB438 84	Homo sapiens	HUMA- Human cancer associated protein sequence SEQ ID NO:1329.	1124	89
1140	gi179281	Homo sapiens	ATP synthase beta subunit precursor (ATPSB) gene, complete cds.	1124	89
1140	gi167413 73	Homo sapiens	Similar to ATP synthase, H+ transporting, mitochondrial F1 complex, beta polypeptide, clone MGC:5231 IMAGE:2900336, mRNA, complete cds.	1124	89
1141	AAW540 79	Homo sapiens	TEXA Homo sapiens BARD1 sequence.	4101	100
1141	gi171017 5	Homo sapiens	Human BRCA1-associated RING domain protein (BARD1) mRNA, complete cds.	4101	100
1141	AAW540 81	Homo sapiens	TEXA Homo sapiens BARD1 P553 sequence.	4097	99
1142	AAW540 89	Homo sapiens	TEXA Homo sapiens BARD1 MR658C sequence.	394	100
1142	AAW540 88	Homo sapiens	TEXA Homo sapiens BARD1 MS761N sequence.	394	100
1142	AAW540	Homo sapiens	TEXA Homo sapiens BARD1	394	100

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
	87		MQ564H sequence.		
1143	gi174322 4	Homo sapiens	MSTP027 (MST027) mRNA, complete cds.	730	100
1143	gi760416 8	Homo sapiens	hypothetical protein FLJ21661, clone MGC:16816 IMAGE:3922036, mRNA, complete cds.	730	100
1143	gi140398 31	Homo sapiens	elongation factor G2 (EFG2) mRNA, complete cds; nuclear gene for mitochondrial product.	730	100
1144	AAM407 29	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 5660.	271	98
1144	AAM389 43	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 2088.	271	98
1144	AAY123 25	Homo sapiens	GEST Human 5' EST secreted protein SEQ ID NO:356.	271	98
1145	gi173903 20	Homo sapiens	clone MGC:9678 IMAGE:3846678, mRNA, complete cds.	872	100
1145	gi120020 02	Homo sapiens	clone 022f05 My030 protein mRNA, complete cds.	872	100
1145	AAY360 68	Homo sapiens	GEST Extended human secreted protein sequence, SEQ ID NO. 453.	844	97
1146	gi173903 20	Homo sapiens	clone MGC:9678 IMAGE:3846678, mRNA, complete cds.	504	95
1146	gi120020 02	Homo sapiens	clone 022f05 My030 protein mRNA, complete cds.	504	95
1146	AAY360 68	Homo sapiens	GEST Extended human secreted protein sequence, SEQ ID NO. 453.	495	94
1147	AAB616 17	Homo sapiens	PROT- Human protein HP10688.	1488	100
1147	gi125784 71	Homo sapiens	unnamed protein product	1488	100
1147	AAY027 81	Homo sapiens	HUMA- Human secreted protein.	1146	77
1148	gi165400 2	Homo sapiens	H.sapiens mRNA for Sop2p-like protein.	572	99
1148	gi128050 63	Mus musculus	actin related protein 2/3 complex, subunit 1A (41 kDa)	567	97
1148	gi126672 58	Rattus norvegicus	suppressor of profilin/p41 of actin-related complex 2/3	567	97
1149	AAB952 58	Homo sapiens	HELI- Human protein sequence SEQ ID NO:17435.	3560	100
1149	gi104347 40	Homo sapiens	cDNA FLJ12957 fis, clone NT2RP2005531, weakly similar to PROTEIN 4.1.	3560	100
1149	gi100471 61	Homo sapiens	mRNA for KIAA1548 protein, partial cds.	2487	100
1150	AAR998 44	Homo sapiens	SUME Human natural killer cell, cell surface mol. NKG7.	552	71
1150	AAQ863 84_aa1	Homo sapiens	ASAN/ G-CSF stimulated human myelocytic cell cDNA.	307	65
1150	AAQ985 51_aa1	Homo sapiens	ASAN/ Granulocyte colony stimulating factor (G-CSF)-	307	65

SEQ ID	Hit ID	Species	Description	S score	Percent identity
			inducible myclocyte gene.		
1151	AAM23781	Homo sapiens	HYSE- Human EST encoded protein SEQ ID NO: 1306.	620	89
1151	AAM23837	Homo sapiens	HYSE- Human EST encoded protein SEQ ID NO: 1362.	481	100
1151	AAY64747	Homo sapiens	GEST Human 5' EST related polypeptide SEQ ID NO:908.	471	98
1152	AAB65277	Homo sapiens	GETH Human PRO1187 (UNQ601) protein sequence SEQ ID NO:399.	656	100
1152	AAY66754	Homo sapiens	GETH Membrane-bound protein PRO1187.	656	100
1152	AAB24077	Homo sapiens	GETH Human PRO1187 protein sequence SEQ ID NO:55.	656	100
1153	gi8331757	Homo sapiens	X28 region near ALD locus containing dual specificity phosphatase 9 (DUSP9), ribosomal protein L18a (RPL18a), Ca2+/Calmodulin-dependent protein kinase I (CAMKI), creatine transporter (CRTR), CDM protein (CDM), adrenoleukodystrophy protein (ALD), plexin-related protein (PLXB3), muscle-specific serine kinase (MSSK), NAD-isocitrate dehydrogenase (IDH), translocon-associated protein delta (TRAP), and LU1 protein (LU1) genes, complete cds; and CCp pseudogene, complete sequence.	1747	92
1153	AAW88436	Homo sapiens	INCY- Disease associated protein kinase DAPK-5.	1667	88
1153	gi3135197	Rattus norvegicus	Ca2+/calmodulin-dependent protein kinase I beta 2	1661	87
1154	AAG67548	Homo sapiens	LEXI- Amino acid sequence of a human transporter protein.	413	98
1154	gi16588684	Homo sapiens	anion transporter/exchanger-8 (SLC26A8) mRNA, complete cds.	413	98
1154	gi15341554	Homo sapiens	putative anion transporter (SLC26A8) mRNA, complete cds.	413	98
1155	AAB64953	Homo sapiens	ROSE/ Human secreted protein sequence encoded by gene 12 SEQ ID NO:131.	995	98
1155	gi12854324	Mus musculus	putative	781	65
1155	gi15217075	Homo sapiens	glioma pathogenesis-related protein (RTVP1) mRNA, complete cds.	443	40
1156	AAB26256	Homo sapiens	UNLO Wild-type human Rac1 protein.	508	100
1156	gi8574038	Homo sapiens	rac1 gene.	508	100
1156	gi2326206	Homo sapiens	Rac3 (RAC3) mRNA, complete cds.	508	100

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
1157	gi104398 53	Homo sapiens	cDNA: FLJ23235 fis, clone CAS04980.	1536	100
1157	AAW855 96	Homo sapiens	MERI Human GABA-A receptor theta subunit.	73	[illegible]
1157	gi159782 47	Yersinia pestis	putative membrane protein	73	[illegible]
1158	gi104398 53	Homo sapiens	cDNA: FLJ23235 fis, clone CAS04980.	1347	90
1158	AAW855 96	Homo sapiens	MERI Human GABA-A receptor theta subunit.	73	35
1158	gi159782 47	Yersinia pestis	putative membrane protein	73	27
1159	gi119331 49	Homo sapiens	mRNA for 6-phosphofructo-2-kinase heart isoform, complete cds.	2452	100
1159	gi309041 9	Homo sapiens	pfkfb2 gene, exons 1 to 15.	2329	99
1159	gi309042 1	Homo sapiens	mRNA for 6-phosphofructo-2-kinase.	2319	98
1160	gi617778 5	Homo sapiens	mRNA for HKR1, partial cds.	3083	99
1160	gi133254 27	Homo sapiens	clone IMAGE:3928207, mRNA, partial cds.	2388	99
1160	gi487783	Homo sapiens	Human zinc finger protein ZNF133.	1592	54
1161	gi145858 59	Homo sapiens	hypothetical protein SB138	1558	98
1161	AAB946 41	Homo sapiens	HELI- Human protein sequence SEQ ID NO:15526.	1543	100
1161	AAG644 03	Homo sapiens	SHAN- Human paneth cell enhanced expression-like protein.	1543	100
1162	AAY536 41	Homo sapiens	CHIR A bone marrow secreted protein designated BMS42.	2182	99
1162	gi966315 3	Homo sapiens	partial mRNA for transport-secretion protein 2.2, (TTS-2.2 gene).	2179	98
1162	gi966315 1	Homo sapiens	partial mRNA for transport-secretion protein 2.1 (TTS-2.1 gene).	2179	98
1163	AAM933 60	Homo sapiens	HELI- Human polypeptide, SEQ ID NO: 2919.	3300	100
1163	gi168782 06	Homo sapiens	hypothetical protein DKFZp434J037, clone MGC:29812 IMAGE:5088037, mRNA, complete cds.	3300	100
1163	gi120532 81	Homo sapiens	mRNA; cDNA DKFZp434J037 (from clone DKFZp434J037); complete cds.	3300	100
1164	AAG812 82	Homo sapiens	ZYMO Human AFP protein sequence SEQ ID NO:82.	3032	100
1164	AAU171 02	Homo sapiens	HUMA- Novel signal transduction pathway protein, Seq ID 667.	3032	100
1164	gi175299 89	Homo sapiens	oxysterol-binding protein-like protein OSBPL9 (OSBPL9) mRNA, complete cds.	3032	100

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
1165	gi14627121	Homo sapiens	Human DNA sequence from clone RP5-824F16 on chromosome 20 Contains the 5' end of the ANGPT4 gene for angiopoietin 4, part of the gene for a novel protein similar to mouse thrombospondin type 1 domain protein R-spondin, ESTs, STSs, GSSs and a CpG island, complete sequence.	230	64
1165	gi16605378	Mus musculus	unnamed protein product	226	42
1165	gi12850680	Mus musculus	putative	226	42
1166	AAF84171_aa1	Homo sapiens	CHUG- Human OATP-B coding sequence.	3573	97
1166	AAZ92403_aa1	Homo sapiens	SCHE cDNA encoding human DC-PGT.	3573	97
1166	AAC61883_aa1	Homo sapiens	CHIR cDNA encoding a human secreted protein.	3573	97
1167	gi30224	Homo sapiens	H.sapiens CRP mRNA for C-reactive protein.	327	100
1167	gi30213	Homo sapiens	H.sapiens mRNA for C-reactive protein.	327	100
1167	gi181068	Homo sapiens	Human C-reactive protein gene, complete cds.	327	100
1168	AAH76194_aa1	Homo sapiens	MILL- Human seven-transmembrane protein 31945 coding sequence.	3429	99
1168	AAB85767	Homo sapiens	MILL- Human seven-transmembrane protein 31945 sequence.	3429	99
1168	gi16551933	Homo sapiens	cDNA FLJ31951 fis, clone NT2RP7007177, weakly similar to Homo sapiens multiple membrane spanning receptor TRC8 mRNA.	3429	99
1169	AAB60299	Homo sapiens	MILL- Human aminopeptidase 17867.	5048	99
1169	AAE04879	Homo sapiens	INCY- Human protease protein-6 (PRTS-6).	5048	99
1169	gi11065900	Homo sapiens	aminopeptidase mRNA, complete cds.	5048	99
1170	gi12844136	Mus musculus	putative	700	52
1170	AAA99905_aa1	Homo sapiens	GETH cDNA encoding human protein PRO846.	662	94
1170	AAB65300	Homo sapiens	GETH Human PRO846 protein sequence SEQ ID NO:517.	662	94
1171	gi12653943	Homo sapiens	clone MGC:2742 IMAGE:2822914, mRNA, complete cds.	3104	100
1171	AAG01236	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 5317.	522	94
1171	gi1304387	Saccharomyces cerevisiae var. diastaticus	glucoamylase	196	22

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
1172	gi13430408	Homo sapiens	BTBD2 protein mRNA, complete cds.	2705	99
1172	gi17026062	Mus musculus	glucose signal repressing protein	1946	77
1172	gi13430406	Homo sapiens	BTBD1 protein mRNA, complete cds.	1937	76
1173	gi17226121	Homo sapiens	F-box protein (FBG4) mRNA, complete cds.	1503	100
1173	gi16553918	Homo sapiens	cDNA FLJ25205 fis, clone REC05844, highly similar to Mus musculus F-box protein FBX17 mRNA.	1503	100
1173	gi15214527	Homo sapiens	Similar to f-box only protein 17, clone MGC:9379 IMAGE:3864760, mRNA, complete cds.	1503	100
1174	AAB88373	Homo sapiens	HELI- Human membrane or secretory protein clone PSEC0109.	2158	100
1174	AAB93215	Homo sapiens	HELI- Human protein sequence SEQ ID NO:12194.	2158	100
1174	AAB93142	Homo sapiens	HELI- Human protein sequence SEQ ID NO:12045.	2158	100
1175	AAB88373	Homo sapiens	HELI- Human membrane or secretory protein clone PSEC0109.	3646	95
1175	AAB93215	Homo sapiens	HELI- Human protein sequence SEQ ID NO:12194.	3646	95
1175	gi14042571	Homo sapiens	cDNA FLJ14791 fis, clone NT2RP4001064, weakly similar to SYNAPTONEMAL COMPLEX PROTEIN SC65.	3646	95
1176	AAB36392	Homo sapiens	CHUG- Human tumour suppressor Gros1-S protein SEQ ID NO:4.	3861	99
1176	gi11127638	Homo sapiens	GROS1-L protein mRNA, complete cds.	3861	99
1176	AAB88373	Homo sapiens	HELI- Human membrane or secretory protein clone PSEC0109.	3847	99
1177	gi10438539	Homo sapiens	cDNA: FLJ22233 fis, clone HRC02016.	2015	100
1177	AAE04892	Homo sapiens	INCY- Human transporter and ion channel-5 (TRICH-5) protein.	2009	99
1177	gi13925661	Mus musculus	sodium/calcium exchanger protein	1708	84 -
1178	AAB65192	Homo sapiens	GETH Human PRO839 (UNQ472) protein sequence SEQ ID NO:167.	366	100
1178	AAG81432	Homo sapiens	ZYMO Human AFP protein sequence SEQ ID NO:382.	366	100
1178	AAY66669	Homo sapiens	GETH Membrane-bound protein PRO839.	366	100
1179	gi599683	Bos taurus	Cleavage and Polyadenylation specificity factor (CPSF) 100kD subunit	4034	98

SEQ ID	Hit ID	Species	Description	S score	Percent identity
1179	gi15489017	Mus musculus	cleavage and polyadenylation specific factor 2, 100kD subunit	3993	97
1179	gi2331036	Mus musculus	cleavage and polyadenylation specificity factor	3993	97
1180	AAZ34648_aa1	Homo sapiens	ZYMO Human growth factor zalpha5 cDNA.	2182	91
1180	AAZ45852_aa1	Homo sapiens	COMP- Human liver angiopoietin-like growth factor DNA sequence.	2182	91
1180	AAA49716_aa1	Homo sapiens	GETH Human PRO179 cDNA clone DNA16451-1078.	2182	91
1181	AAH23183_aa1	Homo sapiens	ISIS- Human macrophage migration inhibitory factor encoding DNA.	564	94
1181	AAB60325	Homo sapiens	KIRI Human wild-type glycosylation-inhibiting factor (GIF).	564	94
1181	AAB85343	Homo sapiens	ISIS- Human macrophage migration inhibitory factor.	564	94
1182	gi7263938	Homo sapiens	mRNA for sodium-glucose cotransporter (SGLT2 gene).	3408	100
1182	gi5679464	Homo sapiens	Human DNA sequence from clone RP1-90G24 on chromosome 22 Contains the RFPL2 gene for RET finger protein-like 2, a Immunoglobulin Lambda Light Chain C region (IGLC) pseudogene, the gene for SAAT1 (low affinity sodium glucosecotransporter (sodium:solute symporter family)) and a Cleavage and Polyadenylation Specific Factor CPSF 160 kD subunit pseudogene. Contains ESTs, GSSs and three putative CpG islands, complete sequence.	3408	100
1182	AAY31221	Homo sapiens	KOEP/ Human SAAT1 protein.	3390	99
1183	gi13623301	Homo sapiens	Similar to Zink transporter 2, clone MGC:11303 IMAGE:3948165, mRNA, complete cds.	1659	100
1183	gi1256378	Rattus norvegicus	zinc transporter ZnT-2	1186	67
1183	gi1763374	Mus musculus	ZnT-3	721	44
1184	AAB92636	Homo sapiens	HELI- Human protein sequence SEQ ID NO:10951.	1480	81
1184	gi7022185	Homo sapiens	cDNA FLJ10260 fis, clone HEMBB1000973, moderately similar to Mus musculus schlafen3 mRNA.	1480	81
1184	AAM40357	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 3502.	1479	81
1185	AAB95351	Homo sapiens	HELI- Human protein sequence SEQ ID NO:17641.	3148	99

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
1185	gi104350 75	Homo sapiens	cDNA FLJ13170 fis, clone NT2RP3003809, weakly similar to SAV PROTEIN.	3148	99
1185	gi126543 21	Homo sapiens	clone MGC:5347 IMAGE:2985725, mRNA, complete cds.	3106	99
1186	gi105675 90	Homo sapiens	sodium bicarbonate cotransporter-like protein mRNA, partial cds.	5645	100
1186	gi134477 47	Homo sapiens	sodium bicarbonate cotransporter NBC4a (NBC4) mRNA, complete cds.	5486	99
1186	gi154195 75	Homo sapiens	sodium bicarbonate cotransporter NBC4c (NBC4) mRNA, complete cds, alternatively spliced.	5382	98
1187	gi101858 24	Homo sapiens	SEBOX (SEBOX) gene, complete cds.	1209	100
1187	gi100921 60	Mus musculus	SEBOX	581	63
1187	gi101858 26	Rattus norvegicus	SEBOX	565	63
1188	AAM435 40	Homo sapiens	HUMA- Human polypeptide SEQ ID NO 218.	246	94
1188	AAM404 68	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 5399.	246	94
1188	AAM386 82	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 1827.	246	94
1189	gi104381 35	Homo sapiens	cDNA: FLJ21924 fis, clone HEP04086.	3703	99
1189	AAM680 38	Homo sapiens	MOLE- Human bone marrow expressed probe encoded protein SEQ ID NO: 28344.	700	100
1189	AAM556 56	Homo sapiens	MOLE- Human brain expressed single exon probe encoded protein SEQ ID NO: 27761.	700	100
1190	gi143885 55	Macaca fascicularis	hypothetical protein	1240	96
1190	AAG932 58	Homo sapiens	NISC- Human protein HP10582.	979	100
1190	gi128548 23	Mus musculus	putative	882	72
1191	AAB945 45	Homo sapiens	HELI- Human protein sequence SEQ ID NO:15296.	2344	99
1191	gi104352 28	Homo sapiens	cDNA FLJ13273 fis, clone OVARC1001010.	2344	99
1191	gi128604 75	Mus musculus	putative	1811	83
1192	gi104404 42	Homo sapiens	mRNA for FLJ00057 protein, partial cds.	3033	99
1192	gi120823 03	Mus musculus	DNA helicase B	1697	61
1192	gi263521 1	Bacillus subtilis	similar to conjugation transfer protein	114	22
1193	AAE1044 5	Homo sapiens	BIOJ Human transmembrane protein (TMP).	2286	100
1193	AAY862	Homo sapiens	HUMA- Human secreted protein	1889	85

SEQ ID	Hit ID	Species	Description	S score	Percent identity
	30		HKFBC53, SEQ ID NO:145.		
1193	AAW68002	Homo sapiens	HUMA- Fragment of human secreted protein encoded by gene 69.	798	96
1194	gi17511729	Homo sapiens	hypothetical protein FLJ12598, clone MGC:31807 IMAGE:4552964, mRNA, complete cds.	1180	100
1194	AAB94162	Homo sapiens	HELI- Human protein sequence SEQ ID NO:14456.	1173	99
1194	gi10434183	Homo sapiens	cDNA FLJ12598 fis, clone NT2RM4001384.	1173	99
1195	gi7243069	Homo sapiens	mRNA for KIAA1344 protein, partial cds.	4137	100
1195	AAB94073	Homo sapiens	HELI- Human protein sequence SEQ ID NO:14262.	1839	99
1195	gi10434025	Homo sapiens	cDNA FLJ12501 fis, clone NT2RM2001681.	1839	99
1196	AAB08894	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 4 SEQ ID NO:51.	235	67
1196	gi335913	Vesicular stomatitis virus	glycoprotein	71	30
1196	gi2960093	*Mycobacterium* tuberculosis H37Rv	*hypothetical protein Rv3669*	71	30
1197	AAM94312	Homo sapiens	HUMA- Human reproductive system related antigen SEQ ID NO: 2970.	1211	98
1197	gi8439396	HERV-H/env62	envelope protein	763	36
1197	gi4959382	Homo sapiens	human endogenous retrovirus HERV-H19 pol protein (pol) gene, partial cds; env protein (env) gene, complete cds; and 3' LTR, complete sequence.	757	36
1198	gi14017381	Homo sapiens	tumor endothelial marker 8 precursor (TEM8) mRNA, complete cds.	1512	100
1198	gi10437939	Homo sapiens	cDNA: FLJ21776 fis, clone HEP00171.	1512	100
1198	gi15987505	Mus musculus	tumor endothelial marker 8 precursor	1484	97
1199	AAB65270	Homo sapiens	GETH Human PRO1158 (UNQ588) protein sequence SEQ ID NO:375.	609	100
1199	AAB87559	Homo sapiens	GETH Human PRO1158.	609	100
1199	AAY66747	Homo sapiens	GETH Membrane-bound protein PRO1158.	609	100
1200	AAM41380	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 6311.	844	87
1200	AAY01785	Homo sapiens	INCY- Human ubiquitin-conjugating enzyme HUBI-1.	818	87
1200	AAY25341	Homo sapiens	PROS- Human NCE-2 protein.	818	87
1201	AAW748	Homo sapiens	HUMA- Human secreted protein	197	100

SEQ ID	Hit ID	Species	Description	S score	Percent identity
	99		encoded by gene 172 clone HODCW06.		
1201	gi137944 93	Guillardia theta	hypothetical protein	67	36
1202	AAS0794 0_aa1	Homo sapiens	AREN- Human cDNA encoding G-protein coupled receptor, hRUP13.	2087	92
1202	AAS1258 3_aa1	Homo sapiens	FARB DNA encoding human serotonin-like G protein-coupled receptor (5-HT-GPCR).	2087	92
1202	AAD195 79_aa1	Homo sapiens	INCY- Human G-protein coupled receptor, GCREC-3 cDNA.	2087	92
1203	AAS0794 0_aa1	Homo sapiens	AREN- Human cDNA encoding G-protein coupled receptor, hRUP13.	2318	100
1203	AAS1258 3_aa1	Homo sapiens	FARB DNA encoding human serotonin-like G protein-coupled receptor (5-HT-GPCR).	2318	100
1203	AAD195 79_aa1	Homo sapiens	INCY- Human G-protein coupled receptor, GCREC-3 cDNA.	2318	100
1204	AAM936 12	Homo sapiens	HELI- Human polypeptide, SEQ ID NO: 3436.	3475	100
1204	gi140431 11	Homo sapiens	Similar to ubiquitin associated and SH3 domain containing, A, clone MGC:15437 IMAGE:2958242, mRNA, complete cds.	3412	100
1204	gi163041 76	Homo sapiens	nm23-phosphorylated unknown substrate mRNA, complete cds.	2759	100
1205	AAB014 24	Homo sapiens	MILL- Human TANGO 213.	1264	100
1205	AAM257 35	Homo sapiens	HYSE- Human protein sequence SEQ ID NO:1250.	1066	100
1205	AAY762 67	Homo sapiens	HUMA- Fragment of human secreted protein encoded by gene 11.	1066	100
1206	AAW749 39	Homo sapiens	HUMA- Human secreted protein encoded by gene 49 clone HAGBI17.	211	100
1207	gi151265 59	Mus musculus	Similar to Cd63 antigen	504	99
1207	gi141980 88	Mus musculus	Cd63 antigen	504	99
1207	gi541060 5	Mus musculus	tetraspanin membrane protein CD63	504	99
1208	AAB951 48	Homo sapiens	HELI- Human protein sequence SEQ ID NO:17173.	3590	99
1208	gi119905 93	Homo sapiens	organic anion transporter polypeptide-related protein 4 (OATPRP4) mRNA, complete cds.	3515	99
1208	AAB491 47	Homo sapiens	BRIM Human organic anion transport protein RP4 protein.	3503	99
1209	AAB733 81	Homo sapiens	NANF- Human gas vesicle protein homologue hGvpT-b.	1866	100

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
1209	gi12005509	Homo sapiens	HT025 mRNA, complete cds.	1866	100
1209	gi10440236	Homo sapiens	cDNA: FLJ23518 fis, clone LNG04878.	1600	100
1210	gi12053021	Homo sapiens	mRNA; cDNA DKFZp434L0714 (from clone DKFZp434L0714); complete cds.	4230	99
1210	AAG64376	Homo sapiens	BIOD- Human II aminoacyl-tRNA synthetase 75.	3517	99
1210	gi10439991	Homo sapiens	cDNA: FLJ23339 fis, clone HEP13401.	3010	99
1211	AAB58941	Homo sapiens	HUMA- Breast and ovarian cancer associated antigen protein sequence SEQ ID 649.	1188	99
1211	gi12005918	Homo sapiens	CDA016 mRNA, complete cds.	1176	98
1211	AAB36615	Homo sapiens	INCY- Human FLEXHT-37 protein sequence SEQ ID NO:37.	862	98
1212	gi10437194	Homo sapiens	cDNA: FLJ21148 fis, clone CAS09413.	2238	96
1212	gi12845648	Mus musculus	putative	1808	80
1212	gi1226033	Saccharomyces cerevisiae	unknown	259	25
1213	gi10437194	Homo sapiens	cDNA: FLJ21148 fis, clone CAS09413.	2203	91
1213	gi12845648	Mus musculus	putative	1774	76
1213	gi1226033.	Saccharomyces cerevisiae	unknown	249	25
1214	gi6630992	Danio rerio	NCC receptor protein 1	418	42
1214	gi12711629	Oreochromis niloticus	nonspecific cytotoxic cell receptor protein	389	41
1214	gi6630998	Ictalurus punctatus	NCC receptor protein 1	361	40
1215	AAC84382_aa1	Homo sapiens	MILL- Human TANGO 209 polypeptide coding sequence.	2430	100
1215	AAS14576_aa1	Homo sapiens	ELIL Human cDNA encoding cysteine-rich secreted protein hCRSP1.	2430	100
1215	AAB48107	Homo sapiens	MILL- Human TANGO 209 polypeptide.	2430	100
1216	ABB11880	Homo sapiens	HYSE- Human lipocortin homologue, SEQ ID NO:2250.	1091	99
1216	AAB43443	Homo sapiens	HUMA- Human cancer associated protein sequence SEQ ID NO:888.	1071	99
1216	AAR22402	Homo sapiens	BIOS Human lipocortin.	1050	99
1217	ABB11880	Homo sapiens	HYSE- Human lipocortin homologue, SEQ ID NO:2250.	1511	100
1217	AAB43443	Homo sapiens	HUMA- Human cancer associated protein sequence SEQ ID NO:888.	1511	100

SEQ ID	Hit ID	Species	Description	S score	Percent identity
1217	AAY084 12	Homo sapiens	UYMC- Human p-40/annexin I protein.	1511	100
1218	ABB1[illegible]88 0	Homo sapiens	HYSE- Human lipocortin homologue, SEQ ID NO:2250.	967	100
1218	AAB4[illegible]4 43	Homo sapiens	HUMA- Human cancer associated protein sequence SEQ ID NO:888.	967	100
1218	AAY084 12	Homo sapiens	UYMC- Human p-40/annexin I protein.	967	100
1219	gi410698 4	Homo sapiens	Human DNA from chromosome 19-specific cosmid R30923, genomic sequence, complete sequence.	2992	100
1219	AAB427 92	Homo sapiens	CURA- Human ORFX ORF2556 polypeptide sequence SEQ ID NO:5112.	2967	99
1219	gi146031 76	Homo sapiens	Similar to RIKEN cDNA 2410153K17 gene, clone MGC:19595 IMAGE:3840843, mRNA, complete cds.	2432	100
1220	AAG814 43	Homo sapiens	ZYMO Human AFP protein sequence SEQ ID NO:404.	492	100
1220	gi150802 20	Homo sapiens	Similar to hypothetical protein, MGC:7764, clone MGC:20548 IMAGE:3607345, mRNA, complete cds.	492	100
1220	gi140361 94	Homo sapiens	unnamed protein product	492	100
1221	AAE0518 3	Homo sapiens	INCY- Human drug metabolising enzyme (DME-14) protein.	2258	100
1221	AAY913 48	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 3 SEQ ID NO:69.	2258	100
1221	gi118545 2	Homo sapiens	Human cytochrome P450 monooxygenase CYP2J2 mRNA, complete cds.	932	44
1222	AAE0518 3	Homo sapiens	INCY- Human drug metabolising enzyme (DME-14) protein.	1344	100
1222	AAY913 48	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 3 SEQ ID NO:69.	1344	100
1222	gi118545 2	Homo sapiens	Human cytochrome P450 monooxygenase CYP2J2 mRNA, complete cds.	548	44
1223	AAW346 18	Homo sapiens	IMUT- Human C3 protein mutant DV-7N.	597	34
1223	AAW346 17	Homo sapiens	IMUT- Human C3 protein mutant DV-6.	597	34
1223	AAW346 16	Homo sapiens	IMUT- Human C3 protein mutant CV-5.	597	34
1224	gi173900 00	Homo sapiens	Similar to RIKEN cDNA 5730455O13 gene, clone MGC:24718 IMAGE:4278022, mRNA, complete cds.	1693	100
1224	AAB417	Homo sapiens	CURA- Human ORFX	1166	100

SEQ ID	Hit ID	Species	Description	S score	Percent identity
	53		ORF1517 polypeptide sequence SEQ ID NO:3034.		
1224	gi128570 19	Mus musculus	putative	1036	87
1225	AAG009 03	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 4984.	294	100
1225	gi339697	Homo sapiens	thymosin beta-10 gene, 3'end.	169	97
1225	gi339687	Homo sapiens	Human thymosin beta-10 mRNA, complete cds.	169	97
1226	AAY913 86	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 41 SEQ ID NO:107.	558	100
1226	gi999289 3	Homo sapiens	phosphoinositol 3-phosphate binding protein-1 (PEPP1) mRNA, complete cds.	75	40
1226	gi101907 44	Homo sapiens] > [Homo sapiens	pleckstrin homology domain-containing, family A (phosphoinositide binding specific) member 4; phosphoinositol 3-phosphate binding protein-1	75	40
1227	AAY600 08	Homo sapiens	META- Human endometrium tumour EST encoded protein 68.	2286	100
1227	AAW747 97	Homo sapiens	HUMA- Human secreted protein encoded by gene 68 clone HKIXR69.	2286	100
1227	gi576230 5	Mus musculus	COP1 protein	2268	99
1228	AAG892 92	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 412.	119	71
1228	AAM937 24	Homo sapiens	HELI- Human polypeptide, SEQ ID NO: 3677.	119	71
1228	gi128036 69	Homo sapiens	CDK4-binding protein p34SEI1, clone MGC:3465 IMAGE:3613213, mRNA, complete cds.	119	71
1229	AAY913 70	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 25 SEQ ID NO:91.	293	100
1230	gi120045 83	Mus musculus	unknown	2566	81
1230	gi128365 62	Mus musculus	putative	2541	80
1230	AAB418 60	Homo sapiens	CURA- Human ORFX ORF1624 polypeptide sequence SEQ ID NO:3248.	1401	100
1231	AAG760 80	Homo sapiens	HUMA- Human colon cancer antigen protein SEQ ID NO:6844.	300	84
1231	AAG013 47	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 5428.	300	84
1231	AAG013 46	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 5427.	300	84
1232	ABB1165 5	Homo sapiens	HYSE- Human secreted protein homologue, SEQ ID NO:2025.	2233	99
1232	gi159297 48	Mus musculus	Unknown (protein for IMAGE:4222865)	1826	81

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
1232	gi128527 13	Mus musculus	putative	1815	81
1233	AAB279 77	Homo sapiens	HUMA- Human secreted protein BLAST search protein SEQ ID NO: 131.	290	96
1233	AAY134 58	Homo sapiens	UYRQ Amino acid sequence of human Fe65.	290	96
1233	gi392493 6	Homo sapiens	Fe65 protein gene, exons 3 through 14 and partial cds.	290	96
1234	AAO077 68	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 21660.	294	100
1234	AAY027 75	Homo sapiens	HUMA- Human secreted protein encoded by gene 12 clone HFTCU19.	288	98
1234	gi729722 6	Drosophila melanogaster	CG4497 gene product	67	42
1235	AAB942 20	Homo sapiens	HELI- Human protein sequence SEQ ID NO:14581.	3191	99
1235	gi104342 88	Homo sapiens	cDNA FLJ12661 fis, clone NT2RM4002189, weakly similar to GLUCOAMYLASE S1/S2 PRECURSOR (EC 3.2.1.3).	3191	99
1235	gi140183 79	Schizosaccharom yces pombe	hypothetical protein; sequence orphan; low similarity to glycoamylases and other cell surface proteins; contains ~250-270 copies of a 13 AA repeat, NSSTPITSSSIL	355	26
1236	AAU035 93	Homo sapiens	INCY- Human DNA modification protein, DNAMP-8.	4977	98
1236	gi606313 7	Mus musculus	F-box protein FBX18	4406	92
1236	AAB942 00	Homo sapiens	HELI- Human protein sequence SEQ ID NO:14538.	4039	99
1237	AAB013 95	Homo sapiens	INCY- Neuron-associated protein.	1507	95
1237	AAU205 29	Homo sapiens	HUMA- Human secreted protein, Seq ID No 521.	1156	86
1237	gi128430 76	Mus musculus	putative	644	95
1238	gi165508 22	Homo sapiens	cDNA FLJ31400 fis, clone NT2NE1000185, weakly similar to UDP-N-ACETYLGLUCOSAMINE--PEPTIDE N-ACETYLGLUCOSAMINYLTRANSFERASE 110 KDA SUBUNIT (EC 2.4.1.-).	1999	96
1238	gi136041 67	Homo sapiens	ARG99 mRNA, complete cds.	781	100
1238	gi672116 1	Arabidopsis thaliana	putative O-linked GlcNAc transferase	372	27
1239	AAY734 14	Homo sapiens	GEMY Human secreted protein clone yb101_1 protein sequence SEQ ID NO:50.	472	100
1239	gi719065	Chlamydia	conserved hypothetical protein	70	28

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
	6	muridarum			
1239	gi332875 9	Chlamydia trachomatis	hypothetical protein	68	28
1240	AAW886 15	Homo sapiens	HUMA- Secreted protein encoded by gene 82 clone HNGBT31.	525	97
1240	AAY840 40	Homo sapiens	USGO Amino acid sequence of cancer associated polypeptide CHI-9a11-2.	71	30
1240	gi148609 75	human herpesvirus 2	DNA polymerase	70	36
1241	gi111214 83	Homo sapiens	mRNA for calsyntenin-2 (CS2 gene).	5080	100
1241	gi111215 06	Mus musculus	calsyntenin-2	4733	94
1241	gi115582 48	Gallus gallus	calsyntenin-1 protein	2962	57
1242	AAM933 76	Homo sapiens	HELI- Human polypeptide, SEQ ID NO: 2951.	1034	100
1242	AAW781 51	Homo sapiens	HUMA- Human secreted protein encoded by gene 26 clone HT3BE24.	1034	100
1242	AAY298 65	Homo sapiens	GEMY Human secreted protein clone pe213_1.	1034	100
1243	gi104395 94	Homo sapiens	cDNA: FLJ23033 fis, clone LNG02005.	2982	99
1243	AAB413 96	Homo sapiens	CURA- Human ORFX ORF1160 polypeptide sequence SEQ ID NO:2320.	2486	100
1243	gi128530 18	Mus musculus	putative	2002	88
1244	AAY122 52	Homo sapiens	GEST Human 5' EST secreted protein SEQ ID NO: 565.	321	92
1244	AAU163 32	Homo sapiens	HUMA- Human novel secreted protein, Seq ID 1285.	320	92
1244	AAU158 73	Homo sapiens	HUMA- Human novel secreted protein, Seq ID 826.	320	92
1245	AAE0487 8	Homo sapiens	INCY- Human protease protein-5 (PRTS-5).	1542	100
1245	gi135436 81	Homo sapiens	clone MGC:14793 IMAGE:4047601, mRNA, complete cds.	1524	99
1245	AAB475 27	Homo sapiens	MILL- Ubiquitin hydrolase-like protein - long form.	1499	100
1246	gi140435 23	Homo sapiens	clone IMAGE:4098694, mRNA, partial cds.	1991	97
1246	gi120608 22	Homo sapiens	serologically defined breast cancer antigen NY-BR-16 mRNA, complete cds.	1991	97
1246	gi129638 69	Mus musculus	gene trap ankyrin repeat containing protein	1980	96
1247	AAB543 57	Homo sapiens	HUMA- Human pancreatic cancer antigen protein sequence SEQ ID NO:809.	301	100
1247	AAY486 00	Homo sapiens	META- Human breast tumour-associated protein 61.	285	98
1247	gi156129	Bacillus	BH0396~unknown conserved	63	32

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
	59	halodurans] > [Bacillus halodurans	protein		
1248	AAM40566	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 5497.	379	46
1248	AAM38780	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 1925.	379	46
1248	gi6539606	Homo sapiens	metastasis suppressor protein mRNA, complete cds.	379	46
1249	gi13992524	Homo sapiens	mRNA for type II alpha phosphatidylinositol 4-kinase gene.	2546	100
1249	gi13111989	Homo sapiens	Similar to hypothetical protein FLJ11105, clone MGC:4395 IMAGE:2905670, mRNA, complete cds.	2546	100
1249	gi13660755	Rattus norvegicus	55 kDa type II phosphatidylinositol 4-kinase	2409	94
1250	AAB95425	Homo sapiens	HELI- Human protein sequence SEQ ID NO:17833.	4871	100
1250	gi10435487	Homo sapiens	cDNA FLJ13465 fis, clone PLACE1003493, weakly similar to ENDOTHELIAL CELL MULTIMERIN PRECURSOR.	4871	100
1250	AAY36300	Homo sapiens	HUMA- Human secreted protein encoded by gene 77.	2472	98
1251	gi6449326	Mus musculus	retinoic acid-responsive protein HA1R-62	83	68
1251	gi6478694	Soybean mosaic virus	P1 protease	67	33
1252	gi12407829	Homo sapiens	netrin 4 precursor (NTN4) mRNA, complete cds.	3361	99
1252	AAG66449	Homo sapiens	GEHO Human beta-netrin.	3347	99
1252	gi11120048	Homo sapiens	beta-netrin mRNA, complete cds.	3347	99
1253	gi16878083	Homo sapiens	enolase 3, (beta, muscle), clone MGC:29581 IMAGE:4902149, mRNA, complete cds.	558	94
1253	gi34789	Homo sapiens	H.sapiens mRNA for muscle specific enolase (MSE) (EC 4.2.1.11).	555	94
1253	gi31170	Homo sapiens	Human ENO3 mRNA for beta-enolase (EC 4.2.1.11).	551	93
1254	AAY07895	Homo sapiens	HUMA- Human secreted protein fragment encoded from gene 44.	537	100
1254	gi17132082	Nostoc sp. PCC 7120	ORF_ID:alr2988~hypothetical protein	69	38
1255	AAB93752	Homo sapiens	HELI- Human protein sequence SEQ ID NO:13419.	1000	99
1255	gi10432785	Homo sapiens	cDNA FLJ11515 fis, clone HEMBA1002241, weakly similar to PROLIFERATING-CELL NUCLEOLAR ANTIGEN P120.	1000	99
1255	gi15680185	Homo sapiens	Similar to RIKEN cDNA 2810405F18 gene, clone	875	99

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
			MGC:22960 IMAGE:4865283, mRNA, complete cds.		
1256	gi14160880	Homo sapiens	PKCI-1-related HIT protein mRNA, complete cds.	827	100
1256	gi13650128	Homo sapiens	HIT-17kDa mRNA, complete cds.	827	100
1256	AAM25739	Homo sapiens	HYSE- Human protein sequence SEQ ID NO:1254.	806	94
1257	gi12849277	Mus musculus	putative	793	93
1257	gi6425295	Caenorhabditis elegans	predicted using Genefinder~contains similarity to Pfam domain: PF00023 (Ank repeat), Score=71.3, E-value=6.5e-18, N=2	200	40
1257	gi4335756	Arabidopsis thaliana	putative ankyrin	195	44
1258	AAB87609	Homo sapiens	GETH Human PRO1890.	1307	99
1258	AAB73309	Homo sapiens	UROG- Human C-type lectin transmembrane antigen PC-LECTIN, SEQ ID NO:2.	1307	99
1258	AAU12441	Homo sapiens	GETH Human PRO1890 polypeptide sequence.	1307	99
1259	AAY05368	Homo sapiens	UYPR- Human HCMV inducible gene protein, SEQ ID NO 4.	1682	97
1259	AAY07036	Homo sapiens	LUDW- Breast cancer associated antigen precursor sequence.	1682	97
1259	gi9956035	Homo sapiens	clone CDABP0047 mRNA sequence.	1682	97
1260	AAB18968	Homo sapiens	INCY- Amino acid sequence of a human transmembrane protein.	1132	100
1260	gi15530313	Homo sapiens	clone MGC:2853 IMAGE:2987806, mRNA, complete cds.	1132	100
1260	gi13937595	Homo sapiens	Similar to RIKEN cDNA 1810017F10 gene, clone MGC:2583 IMAGE:3140820, mRNA, complete cds.	1132	100
1261	AAB18968	Homo sapiens	INCY- Amino acid sequence of a human transmembrane protein.	926	85
1261	gi15530313	Homo sapiens	clone MGC:2853 IMAGE:2987806, mRNA, complete cds.	926	85
1261	gi13937595	Homo sapiens	Similar to RIKEN cDNA 1810017F10 gene, clone MGC:2583 IMAGE:3140820, mRNA, complete cds.	926	85
1262	AAB94434	Homo sapiens	HELI- Human protein sequence SEQ ID NO:15054.	1629	100
1262	gi17512103	Homo sapiens	hypothetical protein FLJ13044, clone MGC:20950 IMAGE:4577143, mRNA, complete cds.	1629	100
1262	gi104348	Homo sapiens	cDNA FLJ13044 fis, clone	1629	100

SEQ ID	Hit ID	Species	Description	S score	Percent identity
	74		NT2RP3001355, weakly similar to TRICARBOXYLATE TRANSPORT PROTEIN PRECURSOR.		
1263	gi178640	Homo sapiens	Human angiotensinogen mRNA, complete CDS.	2366	96
1263	gi119749 7	Homo sapiens	H.sapiens angiotensinogen gene exon 2 (and joined CDS).	2366	96
1263	AAB673 50	Homo sapiens	UTAH Human angiotensinogen protein.	2363	96
1264	gi240126 1	Homo sapiens	HLA-C gene (HLA-Cw*0701 allele), complete cds.	1099	98
1264	gi152772 17	Homo sapiens	genomic DNA, chromosome 6p21.3, HLA Class I region, section 7/20.	1099	98
1264	gi147819 7	Homo sapiens	H.sapiens mRNA for human leukocyte antigen C alpha chain.	1099	98
1265	gi386775	Homo sapiens	Human MHC class I HLA-B8 chain gene (A1,2; B5,8), complete cds.	1033	92
1265	gi240125 9	Homo sapiens	HLA-B gene (HLA-B*0801 allele), complete cds.	1033	92
1265	gi152772 16	Homo sapiens	genomic DNA, chromosome 6p21.3, HLA Class I region, section 6/20.	1033	92
1266	AAM237 60	Homo sapiens	HYSE- Human EST encoded protein SEQ ID NO: 1285.	937	100
1266	gi458664	Homo sapiens	Human MHC class I antigen HLA-B (HLA-B-0704 allele) mRNA, complete cds.	937	100
1266	gi307221	Homo sapiens	Human MHC HLA protein (allele B7) complete cds.	937	100
1267	gi32181	Homo sapiens	H.sapiens HLA-Bw57 gene.	977	88
1267	gi307222	Homo sapiens	Human MHC HLA protein, allele B57, complete cds.	977	88
1267	gi674637 1	Homo sapiens	HLA class I histocompatibility antigen B-57 (HLA-B57) mRNA, complete cds.	971	88
1268	gi149705 74	Homo sapiens	HLA-A gene for MHC class I antigen, allele HLA-A*68011, exons 1-8.	1801	94
1268	gi172403 4	Homo sapiens	Human HLA class I A locus antigen A*68new mRNA, complete cds.	1796	93
1268	gi613877 0	Homo sapiens	HLA-A gene for MHC Class I antigen, A*68 allele, exons 1-8.	1792	93
1269	gi307225	Homo sapiens	Human MHC HLA protein, allele A25, complete cds.	1160	96
1269	gi142503 59	Homo sapiens	clone MGC:17191 IMAGE:4157200, mRNA, complete cds.	1160	96
1269	gi152697 6	Homo sapiens	H.sapiens mRNA for human leucocyte antigen, HLA-A25.	1152	96
1270	gi645336 5	Homo sapiens	mRNA for human leucocyte antigen B (HLA-B gene, B*1501102N allele).	314	88
1270	AAY647	Homo sapiens	GEST Human 5' EST related	148	81

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
	49		polypeptide SEQ ID NO:910.		
1270	AAP70155	Homo sapiens	BEHW Sequence encoded by genomic DNA encoding human histocompatibilityantigen HLA-B 27.	138	62
1271	gi825674	Homo sapiens	H.sapiens gene encoding HLA-Cw6, exons 1-3.	1120	99
1271	gi297097	Homo sapiens	H.sapiens mRNA for HLA-Cw*0602.	1120	99
1271	gi1944480	Homo sapiens	mRNA for HLA-Cw*0602, partial cds.	1120	99
1272	gi2225890	Homo sapiens	Human HLA-A26null allele, complete cds.	977	85
1272	gi487909	Homo sapiens	mRNA for HLA-A11 antigen A11.1, complete cds.	847	94
1272	gi4468256	Homo sapiens	mRNA for MHC class I antigen, allele A*1103.	847	94
1273	gi860968	Homo sapiens	Human HLA-A1 gene.	1122	100
1273	gi8250245	Homo sapiens	HLA-A*0101 gene for MHC class I antigen, exons 1-8.	1122	100
1273	gi386893	Homo sapiens	Human MHC class I HLA-A1 chain gene (A1,2; B8,5), complete cds.	1122	100
1274	AAB94486	Homo sapiens	HELI- Human protein sequence SEQ ID NO:15170.	2982	100
1274	AAM94018	Homo sapiens	HELI- Human stomach cancer expressed polypeptide SEQ ID NO 106.	2982	100
1274	gi14042496	Homo sapiens	cDNA FLJ14750 fis, clone NT2RP3002948, weakly similar to RING CANAL PROTEIN.	2982	100
1275	gi12053277	Homo sapiens	mRNA; cDNA DKFZp434B227 (from clone DKFZp434B227); complete cds.	2242	99
1275	gi10440305	Homo sapiens	cDNA: FLJ23571 fis, clone LNG12303.	2124	94
1275	gi11611603	Macaca fascicularis	hypothetical protein	2064	90
1276	gi8918932	Mus musculus	unnamed protein product	2826	95
1276	gi6329812	Homo sapiens	mRNA for KIAA1130 protein, partial cds.	2716	100
1276	AAS14595_aa1	Homo sapiens	MILL- Human cDNA encoding a novel glycosyltransferase 33877.	1606	58
1277	AAB90676	Homo sapiens	GEMY Human BV141_2 protein sequence SEQ ID 28.	400	98
1277	AAW58985	Homo sapiens	GEMY Homo sapiens adult brain clone BV141_2 encoded protein.	201	100
1277	gi2950481	Schizosaccharomyces pombe	hypothetical protein	71	34
1278	AAY14455	Homo sapiens	HUMA- Human secreted protein encoded by gene 45 clone HCFBJ91.	284	100
1279	AAB85885	Homo sapiens	HELI- Human adenylate kinase 3 (AK3)-like protein.	135	78

SEQ ID	Hit ID	Species	Description	S score	Percent identity
1279	AAB93487	Homo sapiens	HELI- Human protein sequence SEQ ID NO:12786.	135	78
1279	AAB93066	Homo sapiens	HELI- Human protein sequence SEQ ID NO:11883.	135	78
1280	AAB65242	Homo sapiens	GETH Human PRO1291 (UNQ659) protein sequence SEQ ID NO:291.	1378	100
1280	AAB87555	Homo sapiens	GETH Human PRO1291.	1378	100
1280	AAY66719	Homo sapiens	GETH Membrane-bound protein PRO1291.	1378	100
1281	AAB95682	Homo sapiens	HELI- Human protein sequence SEQ ID NO:18481.	1675	96
1281	gi14041989	Homo sapiens	cDNA FLJ14456 fis, clone HEMBB1001915, moderately similar to UBIQUITIN CARBOXYL-TERMINAL HYDROLASE 64E (EC 3.1.2.15).	1675	96
1281	gi7295436	Drosophila melanogaster	Ubp64E gene product	892	71
1282	AAW67841	Homo sapiens	HUMA- Human secreted protein encoded by gene 35 clone HOABG65.	500	100
1282	AAY12238	Homo sapiens	GEST Human 5' EST secreted protein SEQ ID NO: 551.	423	100
1282	AAY11953	Homo sapiens	GEST Human 5' EST secreted protein SEQ ID No: 553.	276	94
1283	AAM25958	Homo sapiens	HYSE- Human protein sequence SEQ ID NO:1473.	484	78
1283	AAW67869	Homo sapiens	HUMA- Human secreted protein encoded by gene 63 clone HHGDB72.	484	78
1283	AAY02285	Homo sapiens	GEMY Secreted protein clone k232_2x polypeptide sequence.	467	78
1284	gi29963	Homo sapiens	Human gene for creatine kinase B (EC 2.7.3.2).	162	71
1284	gi180570	Homo sapiens	Human creatine kinase isozyme CK-B gene, exon 8.	162	71
1284	gi180555	Homo sapiens	Human creatine kinase-B mRNA, complete cds.	162	71
1285	gi12849820	Mus musculus	putative	1170	71
1285	AAM25389	Homo sapiens	HYSE- Human protein sequence SEQ ID NO:904.	790	98
1285	AAY24917	Homo sapiens	INCY- Human phosphatase HPA-2.	550	39
1286	gi12849820	Mus musculus	putative	1456	85
1286	AAY24917	Homo sapiens	INCY- Human phosphatase HPA-2.	798	48
1286	gi8979825	Homo sapiens	Human DNA sequence from clone RP4-776F14 on chromosome 20p12.2-13. Contains the 5' end of the FKBP1A gene for FK506-binding protein 1A (12kD), the	798	48

SEQ ID	Hit ID	Species	Description	S score	Percent identity.
			gene for P47 protein, part of a novel member of the PTPNS (protein tyrosine phosphatase, non-receptor type substrate 1) gene family, ESTs, STSs, GSSs and two CpG islands, complete sequence.		
1287	gi317190 8	Homo sapiens	mRNA for DnaJ protein.	659	100
1287	gi160418 37	Homo sapiens	DnaJ (Hsp40) homolog, subfamily A, member 2, clone MGC:9488 IMAGE:3922477, mRNA, complete cds.	659	100
1287	gi152783 95	Homo sapiens	Similar to DnaJ (Hsp40) homolog, subfamily A, member 2, clone MGC:819 IMAGE:3505399, mRNA, complete cds.	659	100
1288	gi157412 21	Homo sapiens	gene overexpressed in astrocytoma mRNA, complete cds.	3391	99
1288	gi135440 35	Homo sapiens	clone IMAGE:3535476, mRNA, partial cds.	2095	100
1288	gi168781 87	Homo sapiens	Similar to gene overexpressed in astrocytoma, clone MGC:29809 IMAGE:5017710, mRNA, complete cds.	2079	100
1289	AAY927 19	Homo sapiens	GENZ Human polycistin.	20114	99
1289	gi904223	Homo sapiens	polycystic kidney disease 1 protein (PKD1) mRNA, complete cds.	20114	99
1289	AAW238 30	Homo sapiens	DEKR- Human PKD1 protein.	20111	99
1290	AAY559 65	Homo sapiens	SUGE- Full length human ZC4 protein.	1906	100
1290	AAY559 34	Homo sapiens	SUGE- Human ZC4 protein.	1808	100
1290	gi278017 3	Homo sapiens	Human DNA sequence from PAC 82J11 and cosmid U134E6 on chromosome Xq22. Contains NIK like and Thyroxin-binding globulin precursor (T4-binding globulin, TBG) genes, ESTs and STSs.	1588	95
1291	gi104380 63	Homo sapiens	cDNA: FLJ21868 fis, clone HEP02432.	1605	99
1291	gi152774 43	Mus musculus	Unknown (protein for MGC:19083)	1379	84
1291	AAB429 53	Homo sapiens	CURA- Human ORFX ORF2717 polypeptide sequence SEQ ID NO:5434.	522	100
1292	gi795926 3	Homo sapiens	mRNA for KIAA1501 protein, partial cds.	1824	100
1292	ABB1748 8	Homo sapiens	HUMA- Human nervous system related polypeptide SEQ ID NO 6145.	984	100
1292	AAB979	Homo sapiens	SHAN- Human G-protein	965	57

SEQ ID	Hit ID	Species	Description	S score	Percent identity
	11		activating protein 129 SEQ ID NO:2.		
1293	gi104430 48	Homo sapiens	Human DNA sequence from clone RP11-465L10 on chromosome 20. Contains 10 CpG islands, ESTs, STSs and GSSs. Contains the gene for a novel protein similar to Drosophila CG11399, the gene for a novel C2H2 type zinc finger protein similar to chicken FZF-1, a Ferritin light polypeptide (FTL) pseudogene, the MMP9 gene for matrix metalloproteinase 9 (gelatinase B, 92kD gelatinase, 92kD type IV collagenase) (CLG4B), a novel gene, the SLC12A5 gene for solute carrier family 12, (potassium-chloride transporter) member 5 (KIAA1176) and the 3' end of gene KIAA1637, complete sequence.	3567	90
1293	AAH282 26_aa1	Homo sapiens	PFIZ Nucleotide sequence of matrix metalloproteinase-9.	3556	90
1293	AAB204 91	Homo sapiens	SMIK Human matrix metalloproteinase-9 (MMP-9).	3556	90
1294	AAH282 26_aa1	Homo sapiens	PFIZ Nucleotide sequence of matrix metalloproteinase-9.	2375	100
1294	AAB204 91	Homo sapiens	SMIK Human matrix metalloproteinase-9 (MMP-9).	2375	100
1294	AAB846 11	Homo sapiens	PFIZ Amino acid sequence of matrix metalloproteinase-9.	2375	100
1295	AAG040 88	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 8169.	601	91
1295	gi897761	Homo sapiens	H.sapiens mRNA for protein phosphatase 5.	450	92
1295	gi455863 8	Homo sapiens	chromosome 19, BAC 82621 (CIT-B-139a18), complete sequence.	450	92
1296	AAY647 86	Homo sapiens	GEST Human 5' EST related polypeptide SEQ ID NO:947.	200	100
1296	gi162640 94	Sinorhizobium meliloti] > [Sinorhizobium meliloti	HYPOTHETICAL PROTEIN	63	35
1297	gi126980 13	Homo sapiens	mRNA for KIAA1734 protein, partial cds.	3889	100
1297	gi104386 94	Homo sapiens	cDNA: FLJ22346 fis, clone HRC06158.	3877	99
1297	AAB943 54	Homo sapiens	HELI- Human protein sequence SEQ ID NO:14875.	2421	99
1298	gi145756 79	Homo sapiens	hemicentin mRNA, complete cds.	10314	89
1298	gi165519 93	Homo sapiens	cDNA FLJ31995 fis, clone NT2RP7009236, weakly similar to BASEMENT MEMBRANE-SPECIFIC HEPARAN	4274	91

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
			SULFATE PROTEOGLYCAN CORE PROTEIN PRECURSOR.		
1298	gi13872813	Homo sapiens	partial mRNA for fibulin-6 (FIBL-6 gene).	2907	99
1299	gi548084	Rattus norvegicus	olfactory cyclic nucleotide-gated channel	2811	93
1299	gi538129	Rattus norvegicus	cyclic nucleotide gated cation channel	2811	93
1299	gi908824	Bos taurus	alpha subunit of CNG-channel expressed in bovine testis and retinal cone	1576	53
1300	AAB41963	Homo sapiens	CURA- Human ORFX ORF1727 polypeptide sequence SEQ ID NO:3454.	514	100
1300	gi15529703	Homo sapiens	importin 9 mRNA, complete cds.	514	100
1300	gi15186758	Mus musculus	RANBP9 isoform 2	514	100
1301	gi10505349	Homo sapiens	regulator of G-protein signaling (RGS8) mRNA, complete cds.	926	100
1301	gi17382046	Homo sapiens	unnamed protein product	926	100
1301	gi2662057	Rattus norvegicus	RGS8	921	98
1302	AAB95302	Homo sapiens	HELI- Human protein sequence SEQ ID NO:17538.	1242	100
1302	gi10434969	Homo sapiens	cDNA FLJ13105 fis, clone NT2RP3002351, weakly similar to Human mRNA for NAD-dependent methylene tetrahydrofolate dehydrogenase cyclohydrolase (EC 1.5.1.15).	1242	100
1302	gi12834726	Mus musculus	putative	873	94
1303	gi546603	human, tumorous liver, mRNA Partial, 2631 nt]. [Homo sapiens	glutamine synthetase	1787	100
1303	gi17512038	Homo sapiens	clone MGC:20095 IMAGE:3352740, mRNA, complete cds.	1787	100
1303	gi15080157	Homo sapiens	glutamate-ammonia ligase (glutamine synthase), clone MGC:20322 IMAGE:4137547, mRNA, complete cds.	1787	100
1304	gi546603	human, tumorous liver, mRNA Partial, 2631 nt]. [Homo sapiens	glutamine synthetase	432	89
1304	gi31833	Homo sapiens	Human mRNA for glutamine synthetase (E.C. 6.3.1.2).	432	89
1304	gi31831	Homo sapiens	Human rearranged mRNA for glutamine synthase.	432	89
1305	gi16551755	Homo sapiens	cDNA FLJ31807 fis, clone NT2RI2009215, moderately similar to ZINC FINGER	492	54

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
			PROTEIN 165.		
1305	AAM41649	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 6580.	462	51
1305	AAM93917	Homo sapiens	HELI- Human polypeptide, SEQ ID NO: 4075.	462	51
1306	AAU16246	Homo sapiens	HUMA- Human novel secreted protein, Seq ID 1199.	1010	96
1306	gi12832845	Mus musculus	putative	585	83
1306	AAU16240	Homo sapiens	HUMA- Human novel secreted protein, Seq ID 1193.	342	95
1307	gi9798452	Homo sapiens	mRNA for putative capacitative calcium channel (trp7 gene).	4470	100
1307	gi5326854	Mus musculus	receptor-activated calcium channel	4392	98
1307	gi2295903	Homo sapiens	Human putative calcium influx channel (htrp3) mRNA, complete cds.	3529	81
1308	gi6114601	Homo sapiens	mRNA for stromal antigen 3 (STAG3 gene).	281	74
1308	gi3090423	Mus musculus	stag3	203	49
1308	gi13195163	Rattus norvegicus	stromal antigen 3	199	47
1309	gi6114601	Homo sapiens	mRNA for stromal antigen 3 (STAG3 gene).	295	82
1309	gi3090423	Mus musculus	stag3	200	55
1309	gi13195163	Rattus norvegicus	stromal antigen 3	198	54
1310	gi9858562	Homo sapiens	Rh type B glycoprotein (RHBG) mRNA, complete cds.	2176	99
1310	gi15718471	Homo sapiens	Rh type B glycoprotein (RHBG) gene, exons 9, and 10 and complete cds.	2176	99
1310	gi14346006	Pan troglodytes	Rh type B glycoprotein	2161	99
1311	gi7243149	Homo sapiens	mRNA for KIAA1384 protein, partial cds.	3377	100
1311	gi12857673	Mus musculus	putative	2817	98
1311	gi7242973	Homo sapiens	mRNA for KIAA1309 protein, partial cds.	913	33
1312	AAY86297	Homo sapiens	HUMA- Human secreted protein HLDCE79, SEQ ID NO:212.	530	100
1312	AAY21623	Homo sapiens	REGC Ligand binding domain of nuclear receptor hGR.	74	32
1312	AAP80919	Homo sapiens	SALK Sequence of the primary protein sequence of human glucocorticoidreceptor (hGR).	74	32
1313	gi11022690	Homo sapiens	ifp1 mRNA for interferon-responsive finger protein 1 long form, complete cds.	4302	99
1313	AAB95586	Homo sapiens	HELI- Human protein sequence SEQ ID NO:18252.	2612	100
1313	gi14042869	Homo sapiens	cDNA FLJ14970 fis, clone THYRO1000501, weakly	2612	100

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
			similar to 52 KD RO PROTEIN.		
1314	gi178613 72	Homo sapiens	lysyl oxidase-like 4 mRNA, complete cds.	984	93
1314	gi166601 31	Homo sapiens	lysyl oxidase-like 4 (LOXL4) mRNA, complete cds.	984	93
1314	gi146694 71	Homo sapiens	lysyl oxidase-related protein C (LOXC) mRNA, complete cds.	984	93
1315	gi165494 49	Homo sapiens	cDNA FLJ30273 fis, clone BRACE2002685, moderately similar to Homo sapiens androgen-regulated short-chain dehydrogenase/reductase 1 (ARSDR1) mRNA.	1612	98
1315	gi128616 68	Mus musculus	putative	1374	84
1315	gi167406 49	Mus musculus	Similar to RIKEN cDNA A930033N07 gene	1315	81
1316	gi142862 86	Homo sapiens	Similar to hypothetical protein FLJ20515, clone MGC:2696 IMAGE:2820596, mRNA, complete cds.	1006	100
1316	AAY530 23	Homo sapiens	GEMY Human secreted protein clone qf662_3 protein sequence SEQ ID NO:52.	990	99
1316	AAE0483 5	Homo sapiens	SUGE- Human SGP001 phosphatase polypeptide.	931	95
1317	AAN500 69_aa1	Homo sapiens	MITU DNA encoding cardiodilatin in plasmid pHANF48.	771	100
1317	AAW981 93	Homo sapiens	CURA- Human atrial natriuretic peptide prohormone.	771	100
1317	AAP5124 1	Homo sapiens	BIOT- Sequence of pre-pro-atrial natriuretic/vasodilatorpolypeptid e (ANVP).	771	100
1318	AAE0618 3	Homo sapiens	HUMA- Human gene 57 encoded secreted protein fragment, SEQ ID NO:245.	3182	89
1318	AAY872 06	Homo sapiens	HUMA- Human secreted protein sequence SEQ ID NO:245.	3182	89
1318	AAE0609 7	Homo sapiens	HUMA- Human gene 57 encoded secreted protein HRACD80, SEQ ID NO:159.	2906	88
1319	AAU089 95	Homo sapiens	MILL- Human G protein-coupled receptor, GPCR, 45449.	410	96
1319	gi122142 87	Homo sapiens	Human DNA sequence from clone RP3-402H5 on chromosome 6p12.3-21.1 Contains ESTs, STSs and GSSs. Contains the 3' part of a gene for a novel 7 transmembrane receptor of the rhodopsin family and a novel gene, complete sequence.	410	96
1319	gi157973 18	Homo sapiens	unnamed protein product	410	96
1320	gi897761	Homo sapiens	H.sapiens mRNA for protein	387	100

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
			phosphatase 5.		
1320	gi4558638	Homo sapiens	chromosome 19, BAC 82621 (CIT-B-139a18), complete sequence.	387	100
1320	gi12805033	Homo sapiens	protein phosphatase 5, catalytic subunit, clone MGC:5260 IMAGE:3459309, mRNA, complete cds.	387	100
1321	AAB56613	Homo sapiens	ROSE/ Human prostate cancer antigen protein sequence SEQ ID NO:1191.	1174	100
1321	gi182644	Homo sapiens	Human FK506-binding protein 25 (FKBP25) mRNA, complete cds.	1169	100
1321	gi182626	Homo sapiens	Human rapamycin binding protein (FK506) mRNA, complete cds.	1169	100
1322	gi15042691	Homo sapiens	sorting nexin 18 (SNX18) mRNA, complete cds.	2895	100
1322	gi15559064	Mus musculus	SNAG1	2440	86
1322	AAW99023	Homo sapiens	MOUN 17G2 peptide sequence.	1605	95
1323	gi12804803	Homo sapiens	clone MGC:4499 IMAGE:2964565, mRNA, complete cds.	1266	100
1323	gi12654515	Homo sapiens	clone MGC:2827 IMAGE:2964565, mRNA, complete cds.	1266	100
1323	AAB54374	Homo sapiens	HUMA- Human pancreatic cancer antigen protein sequence SEQ ID NO:826.	1261	99
1324	gi11907599	Homo sapiens	protein kinase HIPK2 mRNA, complete cds.	6242	99
1324	AAB65661	Homo sapiens	SUGE- Novel protein kinase, SEQ ID NO: 188.	6086	97
1324	gi17225377	Homo sapiens	homeodomain interacting protein kinase 2 (HIPK2) mRNA, complete cds.	6083	97
1325	AAB65661	Homo sapiens	SUGE- Novel protein kinase, SEQ ID NO: 188.	6124	99
1325	gi17225377	Homo sapiens	homeodomain interacting protein kinase 2 (HIPK2) mRNA, complete cds.	6121	99
1325	gi11907599	Homo sapiens	protein kinase HIPK2 mRNA, complete cds.	6072	97
1326	gi16552298	Homo sapiens	cDNA FLJ32230 fis, clone PLACE6004464, weakly similar to Human placenta (Diff48) mRNA.	3064	99
1326	gi13274202	Homo sapiens	Human DNA sequence from clone RP4-530I15 on chromosome 20 Contains the 3' end of the PTPN1 gene for protein tyrosine phosphatase, non-receptor type 1 (EC 3.1.3.48), the gene for a novel	2261	99

SEQ ID	Hit ID	Species	Description	S score	Percent identity
[redacted]			protein similar to placental protein DIFF40, an RPL36 (60S ribosomal protein L36) pseudogene, a novel gene, two putative novel genes, ESTs, STSs and GSSs, complete sequence.		
1326	gi2224713	Homo sapiens	Human mRNA for KIAA0386 gene, complete cds.	963	34
1327	AAU12177	Homo sapiens	GETH Human PRO305 polypeptide sequence.	1472	82
1327	AAY81487	Homo sapiens	FUJY Human cathepsin L2.	1472	82
1327	AAY02358	Homo sapiens	ONOY Polypeptide identified by the signal sequence trap method.	1472	82
1328	AAU12177	Homo sapiens	GETH Human PRO305 polypeptide sequence.	1698	84
1328	AAY81487	Homo sapiens	FUJY Human cathepsin L2.	1698	84
1328	AAY02358	Homo sapiens	ONOY Polypeptide identified by the signal sequence trap method.	1698	84
1329	AAY87329	Homo sapiens	INCY- Human signal peptide containing protein HSPP-106 SEQ ID NO:106.	692	94
1329	gi15145428	Caenorhabditis elegans	Hypothetical protein Y22D7AL.14	74	23
1329	gi3618016	Human immunodeficiency virus type 1	nef	73	35
1330	AAB82315	Homo sapiens	UYCO Human immunoglobulin receptor isoform IRTA2c.	1120	99
1330	AAB82314	Homo sapiens	UYCO Human immunoglobulin receptor isoform IRTA2b.	1120	99
1330	AAB82313	Homo sapiens	UYCO Human immunoglobulin receptor isoform IRTA2a.	1120	99
1331	gi14336757	Homo sapiens	16p13.3 sequence section 6 of 8.	1178	100
1331	gi13436269	Homo sapiens	hypothetical protein FLJ20898, clone MGC:10688 IMAGE:3622114, mRNA, complete cds.	1178	100
1331	AAG81430	Homo sapiens	ZYMO Human AFP protein sequence SEQ ID NO:378.	954	100
1332	gi5689557	Homo sapiens	mRNA for KIAA1110 protein, partial cds.	3820	99
1332	ABB11713	Homo sapiens	HYSE- Human KIAA1110 protein homologue, SEQ ID NO:2083.	3809	99
1332	gi3882247	Homo sapiens	mRNA for KIAA0763 protein, complete cds.	1587	44
1333	gi13366277	Homo sapiens	Human DNA sequence from clone RP5-998H6 on chromosome 20q13.1. Contains the gene for the ortholog of rat PB-Cadherin, ESTs, STSs, GSS, two CpG islands and genomic marker D20S17, complete	4283	100

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
			sequence.		
1333	gi476057 8	Mus musculus	PB-Cadherin	3918	92
1333	gi139890 6	Rattus norvegicus	long type PB-cadherin	3907	92
1334	gi302304 4	Enterococcus faecalis	orfC	302	45
1334	gi496520	Streptococcus pyogenes	orf iota	271	39
1334	gi129570 24	Enterococcus faecalis	hypothetical protein	251	38
1335	gi128384 81	Mus musculus	putative	491	71
1335	gi207249 7	Mus musculus	perforatorial protein PERF 15	481	68
1335	gi151952 3	Rattus norvegicus	PERF15 protein	477	68
1336	gi104384 54	Homo sapiens	cDNA: FLJ22171 fis, clone HRC00654.	3750	100
1336	gi104381 50	Homo sapiens	cDNA: FLJ21935 fis, clone HEP04373.	3734	99
1336	gi135295 54	Mus musculus	Similar to hypothetical protein FLJ21935	3203	85
1337	gi128034 45	Homo sapiens	clone MGC:2217 IMAGE:3139026, mRNA, complete cds.	463	100
1337	gi729634 4	Drosophila melanogaster	CG4186 gene product	201	47
1337	gi529133	Saccharomyces cerevisiae	Yhr116wp	113	41
1338	gi104343 52	Homo sapiens	cDNA FLJ12697 fis, clone NT2RP1000522, weakly similar to UBIQUITIN CARBOXYL-TERMINAL HYDROLASE DUB-1 (EC 3.1.2.15).	6400	99
1338	AAB951 46	Homo sapiens	HELI- Human protein sequence SEQ ID NO:17169.	6396	99
1338	AAB746 71	Homo sapiens	INCY- Human protease and protease inhibitor PPIM-4.	4021	99
1339	gi108016 26	Macaca fascicularis	hypothetical protein	1668	98
1339	gi128367 18	Mus musculus	putative	1439	84
1339	gi104383 11	Homo sapiens	cDNA: FLJ22054 fis, clone HEP09634.	1351	99
1340	gi175120 67	Homo sapiens	hypothetical protein DKFZp434D0421, clone MGC:20807 IMAGE:4330507, mRNA, complete cds.	1903	100
1340	gi140437 17	Homo sapiens	hypothetical protein DKFZp434D0421, clone MGC:14446 IMAGE:4304040, mRNA, complete cds.	1903	100
1340	gi120531 19	Homo sapiens	mRNA; cDNA DKFZp434D0421 (from clone DKFZp434D0421); complete cds.	1903	100

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
1341	AAB087 65	Homo sapiens	INCY- A human leukocyte and blood related protein (LBAP).	716	93
1341	AAM409 91	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 5922.	508	93
1341	AAB747 18	Homo sapiens	INCY- Human membrane associated protein MEMAP-24.	456	96
1342	AAB955 63	Homo sapiens	HELI- Human protein sequence SEQ ID NO:18202.	3214	99
1342	gi104359 04	Homo sapiens	cDNA FLJ13782 fis, clone PLACE4000489, weakly similar to PROTEIN GRAINY-HEAD.	3214	99
1342	gi128327 62	Mus musculus	putative	2094	94
1343	AAG893 36	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 456.	801	100
1343	AAG813 52	Homo sapiens	ZYMO Human AFP protein sequence SEQ ID NO:222.	801	100
1343	AAY914 23	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 11 SEQ ID NO:144.	801	100
1344	AAY949 78	Homo sapiens	GEMY Human secreted protein clone pw337_6 2nd protein sequence SEQ ID NO:238.	444	100
1344	gi759453 4	Arabidopsis thaliana	putative protein	79	31
1345	gi430989 4	Homo sapiens	PAC clone RP4-555L14 from 7q34-q36, complete sequence.	818	100
1345	gi176464 48	Mus musculus	gammaN-crystallin	724	83
1345	gi176464 46	Homo sapiens	gammaN-crystallin variant (CRYGN) mRNA, complete cds.	600	100
1346	AAY762 16	Homo sapiens	HUMA- Human secreted protein encoded by gene 93.	225	97
1347	AAY114 47	Homo sapiens	GEST Human 5' EST secreted protein SEQ ID No 269.	210	97
1347	gi761976 9	Streptomyces coelicolor A3(2)	probable LacI-family transcriptional regulatory protein.	66	53
1347	gi755580	Streptomyces lividans	ORF-RDR; LacI homolog, similar to E. coli Lac repressor, Swiss-Prot Accession Number P03023	66	53
1348	gi173910 52	Homo sapiens	clone MGC:9915 IMAGE:3871205, mRNA, complete cds.	2220	100
1348	AAG741 53	Homo sapiens	HUMA- Human colon cancer antigen protein SEQ ID NO:4917.	1072	98
1348	gi525318	Haematobia irritans	putative ATPase	937	44
1349	gi173910 52	Homo sapiens	clone MGC:9915 IMAGE:3871205, mRNA, complete cds.	1919	88
1349	AAG741 53	Homo sapiens	HUMA- Human colon cancer antigen protein SEQ ID NO:4917.	782	76
1349	gi525318	Haematobia	putative ATPase	749	39

SEQ ID	Hit ID	Species	Description	S score	Percent identity
		irritans			
1350	AAB43315	Homo sapiens	CURA- Human ORFX ORF3079 polypeptide sequence SEQ ID NO:6158.	1463	99
1350	gi6682590	Homo sapiens	chromosome 19 clones RP11-111O16 and RP11-61F4 containing genes for nuclear receptor coactivator NCoA-62 (nuclear receptor coactivator NCoA-62) gene, complete cds; and unknown gene, complete sequence.	1463	99
1350	AAW82003	Homo sapiens	GEMY Human foetal brain secreted protein fh3_6 (alternative sequence).	1249	99
1351	gi12841145	Mus musculus	putative	1153	93
1351	gi13529212	Homo sapiens	Similar to RIKEN cDNA 1810018M11 gene, clone MGC:12485 IMAGE:3932127, mRNA, complete cds.	1136	99
1351	AAY96202	Homo sapiens	UYNY IkappaB kinase (IKK) binding protein, Y2H56.	1126	98
1352	gi15126788	Mus musculus	Similar to ferritin heavy chain	947	100
1352	gi50954	Mus musculus	ferrerin H subunit	947	100
1352	gi50952	Mus musculus	ferritin heavy subunit (AA 1 - 182)	947	100
1353	AAB70538	Homo sapiens	CURA- Human PRO8 protein sequence SEQ ID NO:16.	2777	98
1353	AAB70537	Homo sapiens	CURA- Human PRO7 protein sequence SEQ ID NO:14.	2777	98
1353	gi13185725	Homo sapiens	n 1755 can be A, G, C, or T.	2777	98
1354	AAB53541	Homo sapiens	HUMA- Human colon cancer antigen protein sequence SEQ ID NO:1081.	110	73
1354	AAR72986	Homo sapiens	GENZ Creatine-kinase subunit B.	110	73
1354	gi29963	Homo sapiens	Human gene for creatine kinase B (EC 2.7.3.2).	110	73
1355	gi12407405	Homo sapiens	tripartite motif protein TRIM9 isoform beta (TRIM9) mRNA, complete cds; alternatively spliced.	2831	100
1355	gi16755524	Rattus norvegicus	Spring	2783	97
1355	gi1665803	Homo sapiens	Human mRNA for KIAA0282 gene, partial cds.	2575	99
1356	AAY53641	Homo sapiens	CHIR A bone marrow secreted protein designated BMS42.	346	98
1356	gi9663153	Homo sapiens	partial mRNA for transport-secretion protein 2.2, (TTS-2.2 gene).	346	98
1356	gi9663151	Homo sapiens	partial mRNA for transport-secretion protein 2.1 (TTS-2.1 gene).	346	98

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
1357	gi108011 97	Homo sapiens	heparanase-like protein HPA2b mRNA, complete cds.	2785	100
1357	gi151326 69	Homo sapiens	unnamed protein product	2785	100
1357	AAA910 97_aa1	Homo sapiens	INSI- Human heparanase, hnhp1, coding sequence.	2626	88
1358	gi633035 8	Homo sapiens	mRNA for KIAA1193 protein, partial cds.	2885	100
1358	AAU162 16	Homo sapiens	HUMA- Human novel secreted protein, Seq ID 1169.	1216	88
1358	AAY158 35	Homo sapiens	PATE/ A human er1 protein.	751	43
1359	AAB749 45	Homo sapiens	YAMA Human ADAM type metal protease MDTS2 protein SEQ ID NO:10.	6065	99
1359	gi114935 89	Homo sapiens	zinc metalloendopeptidase (ADAMTS10) mRNA, partial cds.	5940	99
1359	AAB723 00	Homo sapiens	HIRO/ Human ADAMTS-10 alternative amino acid sequence.	5484	97
1360	gi173842 56	Homo sapiens	partial MUC5AC gene for mucin 5, clone A.	1291	80
1360	gi563375	Homo sapiens	H.sapiens (JER47) MUC5AC mRNA for mucin (partial).	978	91
1360	gi173842 54	Homo sapiens	partial mRNA for mucin 5 (MUC5AC gene).	905	75
1361	AAM243 95	Homo sapiens	HYSE- Human EST encoded protein SEQ ID NO: 1920.	634	100
1361	AAY761 79	Homo sapiens	HUMA- Human secreted protein encoded by gene 56.	634	100
1361	AAB473 27	Homo sapiens	CURA- FCTR4.	74	27
1362	gi150802 64	Homo sapiens	clone MGC:20279 IMAGE:3949150, mRNA, complete cds.	1043	100
1362	gi104390 83	Homo sapiens	cDNA: FLJ22623 fis, clone HSI05687.	1043	100
1362	gi173894 37	Homo sapiens	hypothetical protein FLJ22623, clone MGC:22173 IMAGE:4274089, mRNA, complete cds.	1031	99
1363	AAH787 30_aa1	Homo sapiens	HUMA- Human HIBCJ89 serine/threonine phosphatase cDNA sequence.	1635	99
1363	AAU205 55	Homo sapiens	HUMA- Human secreted protein, Seq ID No 547.	1635	99
1363	AAU206 63	Homo sapiens	HUMA- Human secreted protein, Seq ID No 655.	1635	99
1364	AAB957 00	Homo sapiens	HELI- Human protein sequence SEQ ID NO:18526.	752	100
1364	gi153418 46	Homo sapiens	hypothetical protein FLJ14107, clone MGC:21030 IMAGE:4389733, mRNA, complete cds.	752	100
1364	gi104364 85	Homo sapiens	cDNA FLJ14107 fis, clone MAMMA1001252.	752	100
1365	AAY673	Homo sapiens	GEMY Human secreted protein	367	100

SEQ ID	Hit ID	Species	Description	S score	Percent identity
	16		BL341_4 amino acid sequence.		
1365	AAY08625	Homo sapiens	GEMY Human secreted protein BL341_4.	367	100
1365	AAW42018	Homo sapiens	JACO/ Clone BL341_4 protein.	363	98
1365	AAU19715	Homo sapiens	HUMA- Human novel extracellular matrix protein, Seq ID No 365.	2225	99
1366	gi6136798	Mus musculus	synaptotagmin VIdeltaTM2	2150	96
1366	gi6136796	Mus musculus	synaptotagmin VIdeltaTM1	2150	96
1367	AAU19715	Homo sapiens	HUMA- Human novel extracellular matrix protein, Seq ID No 365.	2173	97
1367	gi6136798	Mus musculus	synaptotagmin VIdeltaTM2	2098	94
1367	gi6136796	Mus musculus	synaptotagmin VIdeltaTM1	2098	94
1368	AAE05175	Homo sapiens	INCY- Human drug metabolising enzyme (DME-6) protein.	2614	97
1368	AAU12225	Homo sapiens	GETH Human PRO4404 polypeptide sequence.	2614	97
1368	gi1193305 6	Sus scrofa	cytochrome P450	1305	50
1369	AAW78135	Homo sapiens	HUMA- Human secreted protein encoded by gene 10 clone HPMGQ80.	385	100
1369	AAO02310	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 16202.	76	39
1369	AAO08772	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 22664.	74	35
1370	gi1004730 1	Homo sapiens	mRNA for KIAA1613 protein, partial cds.	3532	100
1370	AAB36171	Homo sapiens	LEXI- Novel human transporter protein SEQ ID NO: 28.	3412	100
1370	AAB36170	Homo sapiens	LEXI- Novel human transporter protein SEQ ID NO: 26.	3408	99
1371	AAB41673	Homo sapiens	CURA- Human ORFX ORF1437 polypeptide sequence SEQ ID NO:2874.	1221	96
1371	AAB61611	Homo sapiens	PROT- Human protein HP03377.	1220	100
1371	AAE03656	Homo sapiens	INCY- Human extracellular matrix and cell adhesion molecule-20 (XMAD-20).	1220	100
1372	gi5295882	Mus musculus	kinesin like protein 9	3618	88
1372	AAB94768	Homo sapiens	HELI- Human protein sequence SEQ ID NO:15849.	3257	99
1372	gi1043596 8	Homo sapiens	cDNA FLJ13832 fis, clone THYRO1000666, highly similar to Mus musculus mRNA for kinesin like protein 9.	3257	99
1373	AAH25568_aa1	Homo sapiens	CURA- Nucleotide sequence of an interferon omega-1 like	3294	100

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
			protein NOV2.		
1373	AAG67523	Homo sapiens	SMIK Amino acid sequence of a human secreted polypeptide.	3294	100
1373	AAB84469	Homo sapiens	CURA- Amino acid sequence of an interferon omega-1 like protein NOV2.	3294	100
1374	gi5923992	Homo sapiens	Human DNA sequence from clone RP5-1043E3 on chromosome 6p21.1-21.2 Contains part of a novel gene, an transcription factor E2F4 pseudogene, ESTs, STSs and GSSs, complete sequence.	945	99
1374	gi14245713	Giardia intestinalis	kinesin-like protein 9	543	44
1374	gi15022394	Leishmania major	possible kinesin-like protein	531	42
1375	AAG74779	Homo sapiens	HUMA- Human colon cancer antigen protein SEQ ID NO:5543.	1405	99
1375	gi7981261	Homo sapiens	Human DNA sequence from clone RP1-50O24 on chromosome 1p35.1-35.3. Contains the 3' end of the gene for a novel protein similar to C. elegans K07B1.7 (Tr:O01886), the gene for a novel protein (translation of cDNA NT2RM2001100 (Em:AK001211)), the SFN gene for stratifin (14-3-3 protein sigma), the gene for a novel protein with DHHC zinc finger domain, the gene for a novel protein (translation of cDNA KAT07271 (Em:AK000484)) and the gene for B120 (C1orf4) (ARID DNA binding domain containing protein). Contains ESTs, STSs, GSSs and six putative CpG islands, complete sequence.	1122	67
1375	AAG81254	Homo sapiens	ZYMO Human AFP protein sequence SEQ ID NO:26.	1118	72
1376	gi8052237	Homo sapiens	ASCL3 gene, CEGP1 gene, C11orf14 gene, C11orf15 gene, C11orf16 gene and C11orf17 gene.	5605	100
1376	gi8052320	Mus musculus	Cegp1 protein	5054	89
1376	AAG67529	Homo sapiens	SMIK Amino acid sequence of a human secreted polypeptide.	3226	61
1377	gi10437045	Homo sapiens	cDNA: FLJ21044 fis, clone CAE11659.	1663	100
1377	gi4206386	Mus musculus	rig-1 protein	1543	72
1377	AAB57088	Homo sapiens	ROSE/ Human prostate cancer antigen protein sequence SEQ	1518	98

SEQ ID	Hit ID	Species	Description	S score	Percent identity
			ID NO:1666.		
1378	AAY02697	Homo sapiens	HUMA- Human secreted protein encoded by gene 48 clone HTNBR95.	165	100
1379	AAY73386	Homo sapiens	INCY- HTRM clone 3279329 protein sequence.	529	100
1379	AAB63162	Homo sapiens	ROSE/ Human secreted protein sequence encoded by gene 29 SEQ ID NO:88.	363	100
1379	AAB95124	Homo sapiens	HELI- Human protein sequence SEQ ID NO:17122.	83	32
1380	gi10198115	Homo sapiens	2P domain potassium channel TREK2 (KCNK10) mRNA, complete cds.	2760	100
1380	gi8452900	Rattus norvegicus	potassium channel TREK-2	2555	95
1380	gi4584799	Mus musculus	TREK-1 K+ channel subunit	1238	64
1381	gi13276655	Homo sapiens	mRNA; cDNA DKFZp761M0423 (from clone DKFZp761M0423); complete cds.	2626	99
1381	AAE04361	Homo sapiens	INCY- Human kinase (PKIN)-2.	2588	97
1381	gi1836161	Rattus sp.	Ca2+/calmodulin-dependent protein kinase IV kinase isoform; CaM-kinase kinase alpha	2468	93
1382	gi12382781	Homo sapiens	OSBP-related protein 4 mRNA, complete cds.	1124	100
1382	gi13359201	Homo sapiens	mRNA for KIAA1664 protein, partial cds.	1036	100
1382	gi14209840	Homo sapiens	oxysterol binding protein 2 (OSBP2) gene, complete cds.	919	100
1383	gi12805553	Mus musculus	Unknown (protein for MGC:7583)	792	99
1383	gi12858656	Mus musculus	putative	787	98
1383	AAM23865	Homo sapiens	HYSE- Human EST encoded protein SEQ ID NO: 1390.	672	83
1384	gi3002925	Homo sapiens	T cell receptor beta chain (TCRBV13S1-TCRBJ2S1) mRNA, complete cds.	679	73
1384	gi2982508	Homo sapiens	mRNA for TCR beta chain, specific for Mage 3/HLA-A2.	667	71
1384	AAM24051	Homo sapiens	HYSE- Human EST encoded protein SEQ ID NO: 1576.	655	100
1385	gi13879262	Mus musculus	RIKEN cDNA 0610011E17 gene	710	97
1385	gi12850353	Mus musculus	putative	710	97
1385	AAB42905	Homo sapiens	CURA- Human ORFX ORF2669 polypeptide sequence SEQ ID NO:5338.	582	79
1386	AAB95062	Homo sapiens	HELI- Human protein sequence SEQ ID NO:16896.	741	99
1386	AAM241	Homo sapiens	HYSE- Human EST encoded	681	100

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
	27		protein SEQ ID NO: 1652.		
1386	AAY873 28	Homo sapiens	INCY- Human signal peptide containing protein HSPP-105 SEQ ID NO:105.	681	100
1387	AAM235 76	Homo sapiens	HYSE- Human EST encoded protein SEQ ID NO: 1101.	691	100
1387	AAB718 99	Homo sapiens	ZYMO Human zalpha48.	306	68
1387	AAE0658 0	Homo sapiens	SAGA Human protein having hydrophobic domain, HP10786.	306	68
1388	AAB718 63	Homo sapiens	MILL- Human h15571 GPCR.	6511	97
1388	gi159874 91	Homo sapiens	tumor endothelial marker 5 precursor (TEM5) mRNA, complete cds.	6511	97
1388	gi159874 99	Mus musculus	tumor endothelial marker 5 precursor	5693	85
1389	AAE0135 4	Homo sapiens	HUMA- Human gene 3 encoded secreted protein HOHBL42, SEQ ID NO:76.	3747	99
1389	gi431420	Mus musculus	MPS1 protein	2714	77
1389	gi505204 8	Rattus norvegicus	Mpg-1 protein	2672	75
1390	AAM242 00	Homo sapiens	HYSE- Human EST encoded protein SEQ ID NO: 1725.	329	100
1390	AAY195 88	Homo sapiens	HUMA- Amino acid sequence of a human secreted protein.	329	100
1391	AAG742 49	Homo sapiens	HUMA- Human colon cancer antigen protein SEQ ID NO:5013.	888	98
1391	gi128351 18	Mus musculus	putative	884	95
1391	gi730176 3	Drosophila melanogaster	CG11900 gene product	497	59
1392	gi128032 69	Homo sapiens	Similar to CG10641 gene product, clone MGC:3052 IMAGE:3343900, mRNA, complete cds.	701	100
1392	gi104419 42	Homo sapiens	clone PP3051 unknown mRNA.	701	100
1392	AAB954 96	Homo sapiens	HELI- Human protein sequence SEQ ID NO:18039.	698	99
1393	AAY125 12	Homo sapiens	GEST Human 5' EST secreted protein SEQ ID NO:543.	320	98
1394	AAB948 33	Homo sapiens	HELI- Human protein sequence SEQ ID NO:15998.	2735	100
1394	gi104362 42	Homo sapiens	cDNA FLJ13941 fis, clone Y79AA1000850.	2735	100
1394	AAB428 18	Homo sapiens	CURA- Human ORFX ORF2582 polypeptide sequence SEQ ID NO:5164.	1115	99
1395	AAB948 33	Homo sapiens	HELI- Human protein sequence SEQ ID NO:15998.	2713	97
1395	gi104362 42	Homo sapiens	cDNA FLJ13941 fis, clone Y79AA1000850.	2713	97
1395	AAB428 18	Homo sapiens	CURA- Human ORFX ORF2582 polypeptide sequence	1093	94

SEQ ID	Hit ID	Species	Description	S score	Percent identity
			SEQ ID NO:5164.		
1396	gi726392 8	Homo sapiens	Human DNA sequence from clone RP1-61A9 on chromosome 1p35.2-36.13 Contains part of the EPHB2 gene for tyrosine-protein kinase, the gene KIAA0478 for a C2H2 type zinc finger gene, ESTs, STSs, GSSs and three putative CpG Islands, complete sequence.	5326	99
1396	AAE0436 2	Homo sapiens	INCY- Human kinase (PKIN)-3.	5308	99
1396	AAU006 91	Homo sapiens	CURA- Ephrin type-A receptor 8-like protein.	5259	99
1397	gi104376 26	Homo sapiens	cDNA: FLJ21511 fis, clone COL05748.	3713	99
1397	gi167414 00	Mus musculus	Similar to hypothetical protein FLJ21511	3125	82
1397	gi145889 31	Saccharomyces cerevisiae	hypothetical protein	690	29
1398	AAB947 13	Homo sapiens	HELI- Human protein sequence SEQ ID NO:15714.	2869	100
1398	gi104357 45	Homo sapiens	cDNA FLJ13664 fis, clone PLACE1011649.	2869	100
1398	gi168772 91	Homo sapiens	Similar to hypothetical protein 24432, clone MGC:21034 IMAGE:4400396, mRNA, complete cds.	2843	99
1399	AAB829 40	Homo sapiens	UYNY Human androgen receptor trapped protein 5 (ART5).	1429	100
1399	AAB560 85	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 9 SEQ ID NO:179.	1429	100
1399	gi104392 04	Homo sapiens	cDNA: FLJ22709 fis, clone HSI13338.	1429	100
1400	gi104396 25	Homo sapiens	cDNA: FLJ23056 fis, clone LNG03287.	1190	100
1400	gi729573 2	Drosophila melanogaster	ft gene product	185	27
1400	gi157409	Drosophila melanogaster	fat protein	185	27
1401	gi139362 85	Mus musculus	TRH4	1332	61
1401	gi128455 40	Mus musculus	putative	1330	61
1401	AAU007 82	Homo sapiens	INCY- Human apoptosis protein, APOP-2.	1092	65
1402	AAU004 75	Homo sapiens	MILL- Human INTERCEPT 394 alternative form protein.	4272	97
1402	AAU004 73	Homo sapiens	MILL- Human INTERCEPT 394 protein.	4089	99
1402	gi104384 50	Homo sapiens	cDNA: FLJ22169 fis, clone HRC00632.	3505	99
1403	gi107988 04	Homo sapiens	HCMOGT-1 mRNA for sperm antigen, complete cds.	3737	98
1403	ABB1229	Homo sapiens	HYSE- Human secreted protein	2811	100

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
	7		homologue, SEQ ID NO:2667.		
1403	AAM252 55	Homo sapiens	HYSE- Human protein sequence SEQ ID NO:770.	2778	98
1404	AAB954 25	Homo sapiens	HELI- Human protein sequence SEQ ID NO:17833.	4871	100
1404	gi104354 87	Homo sapiens	cDNA FLJ13465 fis, clone PLACE1003493, weakly similar to ENDOTHELIAL CELL MULTIMERIN PRECURSOR.	4871	100
1404	AAY363 00	Homo sapiens	HUMA- Human secreted protein encoded by gene 77.	2472	98
1405	gi100472 11	Homo sapiens	mRNA for KIAA1573 protein, partial cds.	6270	100
1405	gi143883 34	Macaca fascicularis	hypothetical protein	5174	99
1405	AAB958 83	Homo sapiens	HELI- Human protein sequence SEQ ID NO:18994.	3679	100
1406	gi147146 04	Homo sapiens	clone MGC:17248 IMAGE:4215164, mRNA, complete cds.	3291	96
1406	AAW803 18	Homo sapiens	SMIK Neurodegenerative polypeptide HHPDZ65var.	2893	100
1406	gi834683 4	Homo sapiens	mRNA for putative acid-sensing ion channel (ASIC4 gene).	2893	100
1407	AAB656 97	Homo sapiens	SUGE- Novel protein kinase, SEQ ID NO: 225.	1647	100
1407	gi140439 28	Homo sapiens	clone IMAGE:4139786, mRNA, partial cds.	1117	100
1407	AAG024 79	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 6560.	442	100
1408	AAR499 43	Homo sapiens	YAMA/ Human hippocampal cholinergic neurotrophic peptide precursor.	881	89
1408	gi704465	Homo sapiens	H.sapiens mRNA for phosphatidylethanolamine binding protein.	881	89
1408	gi435638	Homo sapiens	Human mRNA for human homologue of rat phosphatidylethanolamine binding protein, complete cds.	881	89
1409	AAB955 17	Homo sapiens	HELI- Human protein sequence SEQ ID NO:18089.	4520	99
1409	gi104357 76	Homo sapiens	cDNA FLJ13687 fis, clone PLACE2000061.	4520	99
1409	gi726465 3	Mus musculus	Kiaa0575	1867	48
1410	gi170283 41	Homo sapiens	hypothetical protein FLJ21820, clone MGC:14932 IMAGE:3611020, mRNA, complete cds.	1732	100
1410	gi104379 97	Homo sapiens	cDNA: FLJ21820 fis, clone HEP01232.	1732	100
1410	gi167697 18	Drosophila melanogaster	LP01162p	437	33
1411	gi150797 29	Homo sapiens	hypothetical protein FLJ21125, clone MGC:14948 IMAGE:4303449, mRNA,	1530	99

SEQ ID	Hit ID	Species	Description	S score	Percent identity
			complete cds.		
1411	gi104371 58	Homo sapiens	cDNA: FLJ21125 fis, clone CAS06077.	1530	99
1411	AAY946 74	Homo sapiens	ZYMO Human zsig83 mature protein sequence.	142	34
1412	gi104381 26	Homo sapiens	cDNA: FLJ21918 fis, clone HEP04006.	3799	100
1412	AAY844 40	Homo sapiens	INCY- Amino acid sequence of a human RNA-associated protein.	2085	59
1412	gi702009 4	Homo sapiens	cDNA FLJ20171 fis, clone COL09761.	1246	66
1413	gi140310 72	Homo sapiens	Human DNA sequence from clone RP3-331H24 on chromosome 6 Contains a putative novel gene, part of the gene for hypothetical protein FLJ21079, similar to opioid growth factor receptor, ESTs, STSs, GSSs and a CpG island, complete sequence.	1307	99
1413	gi104370 94	Homo sapiens	cDNA: FLJ21079 fis, clone CAS02253.	1307	99
1413	gi128434 68	Mus musculus	putative	922	73
1414	AAB943 98	Homo sapiens	HELI- Human protein sequence SEQ ID NO:14968.	3638	100
1414	gi104347 85	Homo sapiens	cDNA FLJ12987 fis, clone NT2RP3000068, weakly similar to SON OF SEVENLESS PROTEIN HOMOLOG 1.	3638	100
1414	AAB956 39	Homo sapiens	HELI- Human protein sequence SEQ ID NO:18376.	3001	100
1415	gi136234 91	Homo sapiens	clone MGC:13125 IMAGE:4111572, mRNA, complete cds.	3054	100
1415	gi165537 93	Homo sapiens	cDNA FLJ25103 fis, clone CBR01405.	1586	74
1415	AAM662 79	Homo sapiens	MOLE- Human bone marrow expressed probe encoded protein SEQ ID NO: 26585.	1301	100
1416	ABB1167 5	Homo sapiens	HYSE- Human secreted protein homologue, SEQ ID NO:2045.	1238	99
1416	gi159289 21	Homo sapiens	hypothetical protein FLJ14393, clone MGC:17935 IMAGE:3916148, mRNA, complete cds.	1238	99
1416	AAY452 72	Homo sapiens	HUMA- Human secreted protein encoded from gene 16.	1236	99
1417	gi136232 49	Homo sapiens	Similar to RIKEN cDNA 3110082I17 gene, clone MGC:11257 IMAGE:3941780, mRNA, complete cds.	945	95
1417	gi128520 07	Mus musculus	putative	466	61
1417	AAW679 36	Homo sapiens	HUMA- Fragment of human secreted protein encoded by	329	98

SEQ ID	Hit ID	Species	Description	S score	Percent identity
			gene 57.		
1418	gi13938274	Homo sapiens	clone MGC:15548 IMAGE:3051320, mRNA, complete cds.	3136	99
1418	AAU16373	Homo sapiens	HUMA- Human novel secreted protein, Seq ID 1326.	963	100
1418	AAU15922	Homo sapiens	HUMA- Human novel secreted protein, Seq ID 875.	963	100
1419	AAB93081	Homo sapiens	HELI- Human protein sequence SEQ ID NO:11914.	1812	98
1419	gi14042419	Homo sapiens	cDNA FLJ14712 fis, clone NT2RP3000825, weakly similar to NEUROGENIC LOCUS NOTCH 3 PROTEIN.	1812	98
1419	AAY72713	Homo sapiens	HUMA- HWAAQ40 clone human attractin-like protein.	1212	99
1420	AAU12418	Homo sapiens	GETH Human PRO1275 polypeptide sequence.	643	98
1420	AAY99379	Homo sapiens	GETH Human PRO1275 (UNQ645) amino acid sequence SEQ ID NO:136.	643	98
1420	AAB25683	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 19 SEQ ID NO:72.	643	98
1421	gi10438712	Homo sapiens	cDNA: FLJ22358 fis, clone HRC06415.	3025	100
1421	gi14211139	Homo sapiens	NADPH oxidase 5 gamma mRNA, complete cds.	3019	99
1421	gi14211137	Homo sapiens	NADPH oxidase 5 alpha mRNA, complete cds.	3019	99
1422	gi12658305	Homo sapiens	kappa B and V(D)J recombination signal sequences binding protein (KRC) mRNA, complete cds.	8934	99
1422	gi10047175	Homo sapiens	mRNA for KIAA1555 protein, partial cds.	8588	99
1422	gi13778886	Mus musculus	DNA binding protein Rc	6216	76
1423	gi17389208	Homo sapiens	clone MGC:16889 IMAGE:3883022, mRNA, complete cds.	2465	100
1423	gi15278167	Homo sapiens	differentiation-related DIF14 long form (DIF14) mRNA, complete cds, alternatively spliced.	2448	99
1423	gi9651220	Mus musculus	LMBR1 long form	2391	96
1424	gi10437078	Homo sapiens	cDNA: FLJ21069 fis, clone CAS01594.	2523	99
1424	gi15929778	Homo sapiens	hypothetical protein FLJ21069, clone MGC:21026 IMAGE:4431888, mRNA, complete cds.	2517	99
1424	gi12859774	Mus musculus	putative	2182	86
1425	AAM93735	Homo sapiens	HELI- Human polypeptide, SEQ ID NO: 3701.	1364	99

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
1425	gi128553 79	Mus musculus	putative	1332	87
1425	gi163593 63	Mus musculus	Similar to hypothetical protein MGC12921	652	53
1426	gi165536 18	Homo sapiens	cDNA FLJ33140 fis, clone UTERU1000160, moderately similar to ZINC FINGER PROTEIN 191.	2173	99
1426	gi104400 85	Homo sapiens	cDNA: FLJ23407 fis, clone HEP19601.	1146	100
1426	gi142501 46	Homo sapiens	hypothetical protein FLJ23407, clone MGC:14819 IMAGE:4248596, mRNA, complete cds.	1143	99
1427	gi127341 04	Homo sapiens	Human DNA sequence from clone RP11-371L19 on chromosome 20. Contains two novel genes, the gene for a novel protein similar to 40S ribosomal protein S10 (RPS10), ESTs, STSs, GSSs and five CpG islands, complete sequence.	2452	100
1427	gi155241 16	Homo sapiens	unnamed protein product	2431	98
1427	gi146024 88	Homo sapiens	clone MGC:10698 IMAGE:3689286, mRNA, complete cds.	2395	98
1428	AAG675 09	Homo sapiens	SMIK Amino acid sequence of a human secreted polypeptide.	4286	100
1428	gi156208 67	Homo sapiens	mRNA for KIAA1904 protein, partial cds.	4272	99
1428	gi319197 5	Homo sapiens	Human DNA sequence from clone RP1-63G5 on chromosome 22q12.3-13.1 Contains the 3' part of the PSCD4 gene for a human SEC7 homolog B2-1 (cytohesin-2, Arno, ARF exchange factor) LIKE protein, a novel gene and the gene coding for a Leucine rich protein. Contains ESTs, STSs, GSSs and three putative CpG islands, complete sequence.	953	100
1429	AAG023 49	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 6430.	468	100
1429	gi177046 2	Homo sapiens	H.sapiens mRNA for M-phase phosphoprotein, mpp6.	468	100
1429	gi150296 28	Homo sapiens	Similar to M-phase phosphoprotein 6, clone MGC:13538 IMAGE:4287267, mRNA, complete cds.	468	100
1430	AAB883 77	Homo sapiens	HELI- Human membrane or secretory protein clone PSEC0113.	239	100
1430	AAB089 04	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 14 SEQ ID NO:61.	239	100
1430	gi142726	Homo sapiens	unnamed protein product	239	100

SEQ ID	Hit ID	Species	Description	S score	Percent identity
	08				
1431	gi9992893	Homo sapiens	phosphoinositol 3-phosphate binding protein-1 (PEPP1) mRNA, complete cds.	3906	95
1431	AAB42086	Homo sapiens	CURA- Human ORFX ORF1850 polypeptide sequence SEQ ID NO:3700.	427	71
1431	gi4589582	Homo sapiens	mRNA for KIAA0969 protein, complete cds.	256	31
1432	gi9992893	Homo sapiens	phosphoinositol 3-phosphate binding protein-1 (PEPP1) mRNA, complete cds.	4152	99
1432	AAB42086	Homo sapiens	CURA- Human ORFX ORF1850 polypeptide sequence SEQ ID NO:3700.	672	99
1432	AAO12592	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 26484.	475	100
1433	AAY94947	Homo sapiens	GEMY Human secreted protein clone cw1292_8 protein sequence SEQ ID NO:100.	408	100
1433	AAB65195	Homo sapiens	GETH Human PRO830 (UNQ470) protein sequence SEQ ID NO:175.	215	64
1433	AAY66672	Homo sapiens	GETH Membrane-bound protein PRO830.	215	64
1434	gi16553818	Homo sapiens	cDNA FLJ25124 fis, clone CBR06414.	1573	100
1434	AAG02137	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 6218.	554	98
1434	AAH75522_aa1	Homo sapiens	SHAN- Human myosin heavy chain 74 encoding cDNA.	275	22
1435	AAF84171_aa1	Homo sapiens	CHUG- Human OATP-B coding sequence.	188	92
1435	AAZ92403_aa1	Homo sapiens	SCHE cDNA encoding human DC-PGT.	188	92
1435	AAC61883_aa1	Homo sapiens	CHIR cDNA encoding a human secreted protein.	188	92
1436	gi11230487	Rattus norvegicus	NTPDase6	501	96
1436	AAB72242	Homo sapiens	HYSE- Mature human CD39 like protein CD39-L2 amino acid sequence.	414	80
1436	AAB72241	Homo sapiens	HYSE- Human CD39 like protein CD39-L2 amino acid sequence.	414	80
1437	gi7243229	Homo sapiens	mRNA for KIAA1424 protein, partial cds.	6604	99
1437	AAB97911	Homo sapiens	SHAN- Human G-protein activating protein 129 SEQ ID NO:2.	6021	99
1437	AAB41660	Homo sapiens	CURA- Human ORFX ORF1424 polypeptide sequence SEQ ID NO:2848.	4377	99
1438	AAB08894	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 4 SEQ ID NO:51.	211	69
1438	gi156262	Buffalopox virus	p8 protein homologue	69	31

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
	57				
1438	gi5830678	variola minor virus	A14L protein	68	27
1439	AAB01713	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 5794.	670	99
1439	gi12001970	Homo sapiens	clone 015h12 My015 protein mRNA, complete cds.	495	96
1439	gi9963910	Xenopus laevis	Churchill protein	495	71
1440	gi14017831	Homo sapiens	mRNA for KIAA1807 protein, partial cds.	1751	100
1440	gi10438885	Homo sapiens	cDNA: FLJ22479 fis, clone HRC10831.	1524	100
1440	gi14424558	Homo sapiens	KIAA0239 protein, clone IMAGE:4301096, mRNA, partial cds.	157	28
1441	AAB43617	Homo sapiens	HUMA- Human cancer associated protein sequence SEQ ID NO:1062.	1481	93
1441	AAP91913	Homo sapiens	BEHW Anticoagulative PP4X.	1481	93
1441	gi189617	Homo sapiens	Human protein PP4-X mRNA, complete cds.	1481	93
1442	AAE03790	Homo sapiens	HUMA- Human gene 9 encoded secreted protein fragment, SEQ ID NO:60.	391	100
1442	AAE03785	Homo sapiens	HUMA- Human gene 9 encoded secreted protein HMWDW68, SEQ ID NO:55.	391	100
1442	AAY73425	Homo sapiens	GEMY Human secreted protein clone yj3_1 protein sequence SEQ ID NO:72.	391	100
1443	gi3882329	Homo sapiens	mRNA for KIAA0804 protein, partial cds.	6282	100
1443	AAB94356	Homo sapiens	HELI- Human protein sequence SEQ ID NO:14879.	4093	99
1443	gi10434628	Homo sapiens	cDNA FLJ12883 fis, clone NT2RP2003981, weakly similar to VACUOLAR PROTEIN SORTING-ASSOCIATED PROTEIN VPS8.	4093	99
1444	gi12539615	Homo sapiens	AKAP-associated sperm protein (ASP) mRNA, complete cds.	1215	99
1444	gi15779077	Homo sapiens	AKAP-associated sperm protein, clone MGC:26950 IMAGE:4820798, mRNA, complete cds.	1212	99
1444	gi11878218	Mus musculus	cAMP-dependent protein kinase regulatory subunit	937	78
1445	AAB43599	Homo sapiens	HUMA- Human cancer associated protein sequence SEQ ID NO:1044.	1283	100
1445	gi14124928	Homo sapiens	clone MGC:3644 IMAGE:2966331, mRNA, complete cds.	1219	100
1445	gi14043853	Homo sapiens	thymidine kinase 1, soluble, clone MGC:14441	1219	100

SEQ ID	Hit ID	Species	Description	S score	Percent identity
			IMAGE:4303880, mRNA, complete cds.		
1446	AAE00404	Homo sapiens	ZYMO Human phosphodiesterase zcytor13 protein.	2733	100
1446	gi13922371	Homo sapiens	unnamed protein product	2733	100
1446	AAM25548	Homo sapiens	HYSE- Human protein sequence SEQ ID NO:1063.	2584	100
1447	gi37664443	Homo sapiens	QDPR gene, exon 1 and joined CDS.	1069	87
1447	gi30819	Homo sapiens	Human mRNA for dihydropteridine reductase (hDHPR).	1069	87
1447	gi181553	Homo sapiens	Human dihydropteridine reductase (hDHPR) mRNA, complete cds.	1069	87
1448	gi13276631	Homo sapiens	mRNA; cDNA DKFZp761F241 (from clone DKFZp761F241); complete cds.	747	100
1448	gi12844872	Mus musculus	putative	650	87
1448	AAY59795	Homo sapiens	META- Human normal ovarian tissue derived protein 72.	554	100
1449	AAB42906	Homo sapiens	CURA- Human ORFX ORF2670 polypeptide sequence SEQ ID NO:5340.	834	100
1449	gi13195151	Homo sapiens	transcription factor TZP (TZP) mRNA, complete cds.	534	54
1449	gi10241461	Homo sapiens	Human DNA sequence from clone RP5-1121G12 on chromosome 20 Contains the 3' end of a gene encoding the hepatocellular carcinoma-associated antigen 58 (HCA58), the SCAND1 gene encoding the domain-containing 1 protein, a novel gene, 2 CpG islands, ESTs, STSs and GSSs, complete sequence.	534	54
1450	AAY12021	Homo sapiens	GEST Human 5' EST secreted protein SEQ ID NO: 334.	265	97
1450	gi9626218	Beet curly top virus	ORF20.1 > [Beet curly top	63	27
1451	AAG01878	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 5959.	348	92
1451	AAB54158	Homo sapiens	HUMA- Human pancreatic cancer antigen protein sequence SEQ ID NO:610.	225	91
1452	AAO08354	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 22246.	1451	89
1452	AAY73384	Homo sapiens	INCY- HTRM clone 2284580 protein sequence.	1451	89
1452	gi13699902	Homo sapiens	mRNA for nucleolar phosphoprotein Nopp34, complete cds.	1451	89
1453	gi142499	Homo sapiens	Similar to hypothetical protein	2387	100

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
	53		FLJ22376, clone MGC:16044 IMAGE:3610443, mRNA, complete cds.		
1453	gi16740559	Homo sapiens	clone MGC:13247 IMAGE:4040497, mRNA, complete cds.	1067	100
1453	gi16551733	Homo sapiens	cDNA FLJ31791 fis, clone NT2RI2008749, weakly similar to SPLICEOSOME ASSOCIATED PROTEIN 49.	1023	53
1454	AAM66321	Homo sapiens	MOLE- Human bone marrow expressed probe encoded protein SEQ ID NO: 26627.	883	50
1454	AAM53933	Homo sapiens	MOLE- Human brain expressed single exon probe encoded protein SEQ ID NO: 26038.	883	50
1454	gi4490388	Felis silvestris	polyprotein	672	44
1455	AAB94815	Homo sapiens	HELI- Human protein sequence SEQ ID NO:15961.	719	100
1455	gi10436205	Homo sapiens	cDNA FLJ13920 fis, clone Y79AA1000420.	719	100
1455	gi12847556	Mus musculus	putative	169	46
1456	gi33044	Homo sapiens	Human mRNA for insulin-like growth factor II (clone P21).	742	97
1456	gi182528	Homo sapiens	Human preproinsulin-like growth factor II (IGF-II) variant mRNA, complete cds.	717	78
1456	AAY70364	Homo sapiens	UYLO- Insulin-like growth factor II.	714	78
1457	AAY99351	Homo sapiens	GETH Human PRO1481 (UNQ750) amino acid sequence SEQ ID NO:41.	1725	100
1457	AAB10259	Homo sapiens	GEMY Human fetal placenta protein fragment BA176_1i.	1631	88
1457	AAB10251	Homo sapiens	GEMY Human adult testes protein fragment AJ142_1i.	761	97
1458	AAB43607	Homo sapiens	HUMA- Human cancer associated protein sequence SEQ ID NO:1052.	664	88
1458	gi695360	Homo sapiens	nuclear-encoded mitochondrial cytochrome c oxidase Va subunit mRNA, complete cds.	658	87
1458	gi12858580	Mus musculus	putative	544	73
1459	gi7549223	Mus musculus	PALS1	3386	96
1459	AAB94180	Homo sapiens	HELI- Human protein sequence SEQ ID NO:14494.	2590	99
1459	gi10434210	Homo sapiens	cDNA FLJ12615 fis, clone NT2RM4001629, weakly similar to MAGUK P55 SUBFAMILY MEMBER 3.	2590	99
1460	gi12697987	Homo sapiens	mRNA for KIAA1721 protein, partial cds.	3859	99
1460	AAB944	Homo sapiens	HELI- Human protein sequence	3853	99

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
	35		SEQ ID NO:15056.		
1460	gi104348 78	Homo sapiens	cDNA FLJ13046 fis, clone NT2RP3001374.	3853	99
1461	gi152144 23	Homo sapiens	clone IMAGE:4563921, mRNA, partial cds.	2603	100
1461	gi179017 49	Homo sapiens	unnamed protein product	2603	100
1461	gi167407 25	Mus musculus	Similar to hexokinase 1	2411	91
1462	AAG755 79	Homo sapiens	HUMA- Human colon cancer antigen protein SEQ ID NO:6343.	714	85
1462	AAB435 66	Homo sapiens	HUMA- Human cancer associated protein sequence SEQ ID NO:1011.	714	85
1462	gi239865 7	Homo sapiens	H.sapiens mRNA translocon-associated protein delta subunit precursor.	714	85
1463	AAG891 28	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 248.	2076	100
1463	gi163068 50	Homo sapiens	hypothetical protein FLJ22637, clone MGC:2443 IMAGE:2821972, mRNA, complete cds.	2076	100
1463	gi104391 04	Homo sapiens	cDNA: FLJ22637 fis, clone HSI06677.	2076	100
1464	AAW642 62	Homo sapiens	BGHM Human neutrophil elastase.	1326	96
1464	AAP8033 5	Homo sapiens	TORA) TORAY IND INC (AOKI/ Sequence of serine protease (SP) of human myeloid cellorigin and leader peptide.	1326	96
1464	gi386981	Homo sapiens	Human neutrophil elastase gene, exon 5.	1326	96
1465	AAY113 85	Homo sapiens	GEST Human 5' EST secreted protein SEQ ID No 207.	220	100
1465	AAB677 83	Homo sapiens	INRM Amino acid sequence of a human thyroid NADPH-oxidase.	75	33
1465	AAM245 05	Homo sapiens	CORI- Colon tumour related amino acid sequence for C799P.	75	33
1466	gi289575 8	Bos taurus	phosphatidic acid-preferring phospholipase A1	4245	91
1466	gi126979 55	Homo sapiens	mRNA for KIAA1705 protein, partial cds.	2582	99
1466	gi165541 84	Homo sapiens	cDNA FLJ25408 fis, clone TST02965, highly similar to Bos taurus phosphatidic acid-preferring phospholipase A1 mRNA.	2378	100
1467	gi147900 25	Homo sapiens	clone MGC:9168 IMAGE:3876839, mRNA, complete cds.	1488	100
1467	gi167686 82	Drosophila melanogaster	HL02815p	1155	49
1467	gi107269 44	Drosophila melanogaster	CG11306 gene product	1155	49
1468	AAB747	Homo sapiens	HUMA- Human secreted protein	902	91

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
	60		sequence encoded by gene 18 SEQ ID NO:69.		
1468	AAB747 59	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 18 SEQ ID NO:68.	902	91
1468	AAB747 50	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 18 SEQ ID NO:59.	902	91
1469	AAB747 60	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 18 SEQ ID NO:69.	1006	96
1469	AAB747 59	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 18 SEQ ID NO:68.	1006	96
1469	AAB747 50	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 18 SEQ ID NO:59.	1006	96
1470	AAW131 08	Homo sapiens	ONYX- Human 14-3-3 beta or HS1 1054.	1225	95
1470	gi279155 2	Homo sapiens	Human DNA sequence from clone RP1-148E22 on chromosome 20q12-13.12 Contains the YWHAB gene encoding tyrosine 3-monooxygenase/ntryptophan 5-monooxygenase activation protein, beta polypeptide, a novel gene similar to PABPC1 (poly (A)-binding protein, cytoplasmic 1), 2 CpG islands, ESTs, STSs and GSSs, complete sequence.	1225	95
1470	gi23114	Homo sapiens	H.sapiens mRNA for HS1 protein.	1225	95
1471	AAB948 34	Homo sapiens	HELI- Human protein sequence SEQ ID NO:16000.	2638	100
1471	gi104362 44	Homo sapiens	cDNA FLJ13942 fis, clone Y79AA1000962, weakly similar to MYOSIN HEAVY CHAIN, NON-MUSCLE.	2638	100
1471	gi142905 66	Homo sapiens	hypothetical protein FLJ13942, clone MGC:9884 IMAGE:3867690, mRNA, complete cds.	1501	100
1472	AAB105 50	Homo sapiens	HOFM/ Human aspartate protease psl 4 protein.	1925	100
1472	AAB088 60	Homo sapiens	INCY- Amino acid sequence of a human secretory protein.	1925	100
1472	AAB089 71	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 25 SEQ ID NO:128.	1920	99
1473	AAG892 62	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 382.	231	100
1473	AAY307 21	Homo sapiens	HUMA- Amino acid sequence of a human secreted protein.	231	100
1473	AAB236 15	Homo sapiens	ALPH- Human secreted protein SEQ ID NO: 30.	222	97
1474	gi702033	Homo sapiens	cDNA FLJ20318 fis, clone	2962	100

SEQ ID	Hit ID	Species	Description	S score	Percent identity
	6		HEP08704.		
1474	AAM40651	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 5582.	1804	47
1474	AAM38865	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 2010.	1804	47
1475	gi14495621	Homo sapiens	hypothetical protein FLJ22578, clone MGC:14892 IMAGE:3506508, mRNA, complete cds.	816	100
1475	gi10439014	Homo sapiens	cDNA: FLJ22578 fis, clone HSI02546.	802	100
1475	AAM72825	Homo sapiens	MOLE- Human bone marrow expressed probe encoded protein SEQ ID NO: 33131.	616	100
1476	gi2224697	Homo sapiens	Human mRNA for KIAA0378 gene, partial cds.	4017	100
1476	gi6681583	Homo sapiens	ELKS mRNA, complete cds.	3463	72
1476	gi13445784	Mus musculus	Rab6-interacting protein 2 isoform A	3423	70
1477	gi15530323	Homo sapiens	clone MGC:4131 IMAGE:2961417, mRNA, complete cds.	3200	99
1477	gi16307502	Mus musculus	Unknown (protein for MGC:11530)	3076	95
1477	gi15277234	Homo sapiens	genomic DNA, chromosome 6p21.3, HLA Class I region, section 12/20.	2227	99
1478	AAY11794	Homo sapiens	GEST Human 5' EST secreted protein SEQ ID No: 394.	375	100
1478	AAB94977	Homo sapiens	HELI- Human protein sequence SEQ ID NO:16558.	68	35
1478	gi10433141	Homo sapiens	cDNA FLJ11800 fis, clone HEMBA1006252.	68	35
1479	AAB08732	Homo sapiens	UYCO Amino acid sequence of a human OLD-35 polypeptide.	3498	98
1479	gi12835817	Mus musculus	putative	2439	89
1479	AAB92684	Homo sapiens	HELI- Human protein sequence SEQ ID NO:11065.	2369	99
1480	gi14424568	Homo sapiens	Mov10 (Moloney leukemia virus 10, mouse) homolog, clone MGC:15000 IMAGE:4109453, mRNA, complete cds.	4983	100
1480	gi12803447	Homo sapiens	Similar to Moloney leukemia virus 10, clone MGC:2948 IMAGE:3138543, mRNA, complete cds.	4983	100
1480	gi10047339	Homo sapiens	mRNA for KIAA1631 protein, partial cds.	4983	100
1481	AAU05584	Homo sapiens	OXFO- Human breast cancer membrane protein 81, BCMP-81.	718	100
1481	AAU25727	Homo sapiens	OXFO- Breast cancer-associated membrane protein (BCMP) 81.	718	100
1481	AAW85738	Homo sapiens	SAGA Polypeptide with transmembrane domain.	718	100

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
1482	gi140179 57	Homo sapiens	mRNA for KIAA1870 protein, partial cds.	1496	94
1482	AAB420 00	Homo sapiens	CURA- Human ORFX ORF1764 polypeptide sequence SEQ ID NO:3528.	1302	93
1482	AAB938 66	Homo sapiens	HELI- Human protein sequence SEQ ID NO:13743.	860	100
1483	gi140179 57	Homo sapiens	mRNA for KIAA1870 protein, partial cds.	1608	100
1483	AAB420 00	Homo sapiens	CURA- Human ORFX ORF1764 polypeptide sequence SEQ ID NO:3528.	1414	100
1483	AAB938 66	Homo sapiens	HELI- Human protein sequence SEQ ID NO:13743.	860	100
1484	gi499184	Felis catus	neuronal protein	617	96
1484	AAB950 41	Homo sapiens	HELI- Human protein sequence SEQ ID NO:16804.	507	77
1484	gi795925 1	Homo sapiens	mRNA for KIAA1495 protein, partial cds.	507	77
1485	AAB604 57	Homo sapiens	INCY- Human cell cycle and proliferation protein CCYPR-5, SEQ ID NO:5.	928	100
1485	gi150805 50	Homo sapiens	hypothetical protein FLJ23467, clone MGC:21000 IMAGE:4509736, mRNA, complete cds.	928	100
1485	gi104401 66	Homo sapiens	cDNA: FLJ23467 fis, clone HSI11213.	925	99
1486	AAB933 01	Homo sapiens	HELI- Human protein sequence SEQ ID NO:12369.	4341	99
1486	gi140426 07	Homo sapiens	cDNA FLJ14812 fis, clone NT2RP4002081, weakly similar to TRANSCRIPTION INITIATION FACTOR IIA ALPHA AND BETA CHAINS.	4341	99
1486	gi100471 79	Homo sapiens	mRNA for KIAA1557 protein, partial cds.	4168	99
1487	AAB948 04	Homo sapiens	HELI- Human protein sequence SEQ ID NO:15939.	2050	100
1487	gi150825 06	Homo sapiens	hypothetical protein FLJ13910, clone MGC:20406 IMAGE:4636136, mRNA, complete cds.	2050	100
1487	gi104361 89	Homo sapiens	cDNA FLJ13910 fis, clone Y79AA1000131.	2050	100
1488	gi128460 13	Mus musculus	putative	1876	97
1488	gi783955 9	Homo sapiens	PAD mRNA, complete cds.	1789	98
1488	gi136041 69	Homo sapiens	ARG147 mRNA, complete cds.	1575	99
1489	gi126979 35	Homo sapiens	mRNA for KIAA1695 protein, partial cds.	2124	100
1489	gi104386 24	Homo sapiens	cDNA: FLJ22297 fis, clone HRC04521.	2124	100
1489	AAB424 21	Homo sapiens	CURA- Human ORFX ORF2185 polypeptide sequence	1571	100

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
			SEQ ID NO:4370.		
1490	AAB475 62	Homo sapiens	INCY- Protease PRTS-4.	4321	99
1490	AAM937 85	Homo sapiens	HELI- Human polypeptide, SEQ ID NO: 3804.	4317	99
1490	gi128363 32	Mus musculus	putative	4152	95
1491	gi165538 16	Homo sapiens	cDNA FLJ25123 fis, clone CBR06154.	1752	93
1491	AAO118 34	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 25726.	1347	98
1491	AAM257 94	Homo sapiens	HYSE- Human protein sequence SEQ ID NO:1309.	919	99
1492	gi140433 13	Homo sapiens	clone IMAGE:3609599, mRNA, partial cds.	780	100
1492	AAY122 25	Homo sapiens	GEST Human 5' EST secreted protein SEQ ID NO: 538.	511	97
1492	AAG005 45	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 4626.	500	97
1493	AAM934 50	Homo sapiens	HELI- Human polypeptide, SEQ ID NO: 3100.	2693	99
1493	AAY077 54	Homo sapiens	HUMA- Human secreted protein fragment encoded from gene 11.	1723	100
1493	AAW790 94	Homo sapiens	GEMY Human secreted protein do568_11.	1699	98
1494	AAG648 94	Homo sapiens	BIOD- Human phosphoenol pyruvate carboxylase 81.	3851	100
1494	AAB952 50	Homo sapiens	HELI- Human protein sequence SEQ ID NO:17415.	1589	100
1494	gi132766 65	Homo sapiens	mRNA; cDNA DKFZp761K1524 (from clone DKFZp761K1524); complete cds.	1493	100
1495	AAB437 37	Homo sapiens	HUMA- Human cancer associated protein sequence SEQ ID NO:1182.	817	89
1495	AAR592 88	Homo sapiens	SHIO Human reg protein.	817	89
1495	gi576455 5	Homo sapiens	lithostathine (REG1A) mRNA, complete cds.	817	89
1496	gi126525 61	Homo sapiens	Similar to cytochrome b-561, clone MGC:3308 IMAGE:3509626, mRNA, complete cds.	1129	96
1496	gi128042 35	Homo sapiens	Similar to cytochrome b-561, clone MGC:2190 IMAGE:3535771, mRNA, complete cds.	1126	95
1496	gi939707	Homo sapiens	Human cytochrome b561 gene, exon 5 and complete cds.	1124	95
1497	gi104370 90	Homo sapiens	cDNA: FLJ21077 fis, clone CAS02152.	2182	99
1497	gi104372 11	Homo sapiens	cDNA: FLJ21159 fis, clone CAS09969.	1885	100
1497	AAB639 60	Homo sapiens	LUDW- Human prostate cancer associated antigen protein sequence SEQ ID NO:1322.	904	96

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
1498	AAB40996	Homo sapiens	CURA- Human ORFX ORF760 polypeptide sequence SEQ ID NO:1520.	3391	95
1498	gi10436963	Homo sapiens	cDNA: FLJ20986 fis, clone CAE01156.	3137	99
1498	AAM93525	Homo sapiens	HELI- Human polypeptide, SEQ ID NO: 3259.	2627	99
1499	AAB99891	Homo sapiens	CHUG- Human RNA helicase gene helicain B protein sequence SEQ ID NO:4.	3818	100
1499	gi16566550	Homo sapiens	DEAD/DEXH helicase DDX31 mRNA, complete cds.	3817	99
1499	gi15215273	Homo sapiens	hypothetical protein FLJ13633, clone MGC:14872 IMAGE:3941452, mRNA, complete cds.	3455	100
1500	gi12005511	Homo sapiens	HT027 mRNA, complete cds.	744	100
1500	gi10436844	Homo sapiens	cDNA: FLJ20886 fis, clone ADKA03257.	739	99
1500	gi11036973	Homo sapiens	HSP22-like protein interacting protein 17 mRNA, complete cds.	459	100
1501	AAB95261	Homo sapiens	HELI- Human protein sequence SEQ ID NO:17444.	906	100
1501	gi10434755	Homo sapiens	cDNA FLJ12967 fis, clone NT2RP2005806.	906	100
1501	AAB94267	Homo sapiens	HELI- Human protein sequence SEQ ID NO:14682.	587	95
1502	gi16553461	Homo sapiens	cDNA FLJ33132 fis, clone UMVEN2000133, weakly similar to RABPHILIN-3A.	2594	99
1502	gi10438690	Homo sapiens	cDNA: FLJ22344 fis, clone HRC06080.	1661	99
1502	AAB93562	Homo sapiens	HELI- Human protein sequence SEQ ID NO:12957.	1289	55
1503	gi16549330	Homo sapiens	cDNA FLJ30165 fis, clone BRACE2000698, weakly similar to ANKYRIN 2.	2280	98
1503	gi12652741	Homo sapiens	clone MGC:3130 IMAGE:3352851, mRNA, complete cds.	2267	78
1503	AAO01850	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 15742.	1975	89
1504	AAB66295	Homo sapiens	ICOS- Human tankyrase2 TANK2-SHORT SEQ ID NO: 135.	6088	99
1504	AAB66294	Homo sapiens	ICOS- Human tankyrase2 TANK2-LONG SEQ ID NO: 133.	6088	99
1504	AAB66290	Homo sapiens	ICOS- Human tankyrase2 clone consensus protein SEQ ID NO: 107.	6088	99
1505	gi5689427	Homo sapiens	mRNA for KIAA1045 protein, partial cds.	2087	99
1505	gi13358652	Macaca fascicularis	hypothetical protein	1205	96
1505	ABB1680	Homo sapiens	HUMA- Human nervous system	370	100

SEQ ID	Hit ID	Species	Description	S score	Percent identity
	6		related polypeptide SEQ ID NO 5463.		
1506	gi568942 7	Homo sapiens	mRNA for KIAA1045 protein, partial cds.	2052	94
1506	gi133586 52	Macaca fascicularis	hypothetical protein	1205	96
1506	ABB1680 6	Homo sapiens	HUMA- Human nervous system related polypeptide SEQ ID NO 5463.	370	100
1507	gi104390 66	Homo sapiens	cDNA: FLJ22612 fis, clone HSI04965.	2767	100
1507	gi145298 86	Mus musculus	bM145O4.1 (novel protein)	2276	78
1507	gi128553 00	Mus musculus	putative	2276	78
1508	gi114933 65	Homo sapiens	Human DNA sequence from clone RP5-1009E24 on chromosome 20 Contains the SN gene encoding sialoadhesin, a novel gene similar to KIAA0417, the CENPB gene for centromere protein B, the CDC25B gene for Cell division cycle protein 25B, three novel genes, the 5' end of gene KIAA1271, nine CpG islands, ESTs, STSs and GSSs, complete sequence.	6334	99
1508	gi126561 30	Homo sapiens	sialoadhesin mRNA, complete cds.	6330	99
1508	gi104404 38	Homo sapiens	mRNA for FLJ00055 protein, partial cds.	5046	99
1509	AAY765 39	Homo sapiens	META- Human ovarian tumor EST fragment encoded protein 35.	261	98
1510	AAE0488 4	Homo sapiens	INCY- Human protease protein-11 (PRTS-11).	424	100
1510	AAB732 63	Homo sapiens	UYAL- Human triacylglycerol hydrolase, TGH.	209	51
1510	gi180950	Homo sapiens	Human carboxylesterase mRNA, complete cds.	209	51
1511	AAB944 05	Homo sapiens	HELI- Human protein sequence SEQ ID NO:14983.	4412	100
1511	gi104347 98	Homo sapiens	cDNA FLJ12994 fis, clone NT2RP3000207, weakly similar to GLUCOAMYLASE S1/S2 PRECURSOR (EC 3.2.1.3).	4412	100
1511	gi104371 27	Homo sapiens	cDNA: FLJ21104 fis, clone CAS04958.	978	100
1512	AAB621 75	Homo sapiens	PLAC Human p110FYVE protein.	4028	99
1512	AAF5740 3_aa1	Homo sapiens	PLAC Human p110FYVE protein encoding DNA.	4027	99
1512	gi113449 51	Homo sapiens	FYVE-finger-containing Rab5 effector protein Rabenosyn-5 mRNA, complete cds.	4027	99
1513	gi795926	Homo sapiens	mRNA for KIAA1500 protein,	4061	100

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
	1		partial cds.		
1513	gi9837427	Lytechinus variegatus	embryonic blastocoelar extracellular matrix protein precursor	1085	34
1513	AAG73354	Homo sapiens	HUMA- Human gene 9-encoded secreted protein HETAM53, SEQ ID NO:125.	517	100
1514	gi16550108	Homo sapiens	cDNA FLJ30829 fis, clone FEBRA2001790, highly similar to Xenopus laevis RRM-containing protein SEB-4 mRNA.	914	100
1514	gi13624461	Homo sapiens	Human DNA sequence from clone RP1-259A10 on chromosome 6p22.1-23 Contains the gene for an ssDNA binding protein (SEB4D), ESTs, STSs, GSSs and a CpG island, complete sequence.	914	100
1514	gi8895698	Xenopus laevis	RRM-containing protein SEB-4	790	88
1515	gi16551580	Homo sapiens	cDNA FLJ31673 fis, clone NT2RJ2005061.	3158	99
1515	AAB94929	Homo sapiens	HELI- Human protein sequence SEQ ID NO:16399.	2275	100
1515	gi10432847	Homo sapiens	cDNA FLJ11565 fis, clone HEMBA1003229.	2275	100
1516	gi3947688	Homo sapiens	mRNA for Sec24 protein (Sec24A isoform), partial.	5355	98
1516	AAM79111	Homo sapiens	HYSE- Human protein SEQ ID NO 1773.	3090	55
1516	gi3947690	Homo sapiens	mRNA for Sec24 protein (Sec24B isoform).	3090	55
1517	AAY12049	Homo sapiens	GEST Human 5' EST secreted protein SEQ ID NO: 362.	253	92
1517	gi4155063	Helicobacter pylori J99	putative	75	29
1517	gi5738522	Schizosaccharomyces pombe	putative pre-mrna splicing factor atp-dependent rna helicase	69	33
1518	AAB20349	Homo sapiens	UYRQ Human vomeronasal-like receptor hVLR1 (long form).	1859	99
1518	AAG64295	Homo sapiens	HELI- Human GTP-binding protein-coupled receptor GPRv31.	1859	99
1518	gi9988585	Homo sapiens	putative pheromone receptor V1RL1 long form (V1RL1) mRNA, complete cds.	1859	99
1519	gi14249909	Homo sapiens	clone IMAGE:3506174, mRNA, partial cds.	2759	90
1519	gi14249907	Homo sapiens	clone IMAGE:3506145, mRNA, partial cds.	2759	90
1519	AAY99355	Homo sapiens	GETH Human PRO1295 (UNQ664) amino acid sequence SEQ ID NO:54.	1265	100
1520	gi16041686	Homo sapiens	hypothetical protein FLJ22393, clone MGC:16798 IMAGE:3916157, mRNA,	1470	100

SEQ ID	Hit ID	Species	Description	S score	Percent identity
			complete cds.		
1520	gi|10438763	Homo sapiens	cDNA: FLJ22393 fis, clone HRC07880.	1463	99
1520	gi|12853419	Mus musculus	putative	1446	98
1521	AAR35072	Homo sapiens	UYPR- Human t-complex associated testes expressed protein 1.	2576	97
1521	gi|201910	Mus musculus	Tctc-1 peptide	1883	74
1521	gi|7300285	Drosophila melanogaster	CG14325 gene product	348	27
1522	gi|12851762	Mus musculus	putative	689	88
1522	AAG02298	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 6379.	382	100
1522	gi|16769034	Drosophila melanogaster	LD15209p	294	38
1523	AAQ90526_aa1	Homo sapiens	OKLA- Human SIII 15 kDa subunit cDNA.	426	100
1523	AAW13850	Homo sapiens	OKLA- Human RNA polymerase transcription factor elongin 15 kDa subunit.	426	100
1523	AAR75087	Homo sapiens	OKLA- Human SIII 15 kDa subunit.	426	100
1524	gi|12855672	Mus musculus	putative	2165	85
1524	AAU17429	Homo sapiens	HUMA- Novel signal transduction pathway protein, Seq ID 994.	987	98
1524	AAG04081	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 8162.	507	99
1525	AAW13085	Homo sapiens	SAGA Human E2 ubiquinone binding enzyme.	667	88
1525	gi|13097195	Homo sapiens	ubiquitin-conjugating enzyme E2N (homologous to yeast UBC13), clone MGC:5063 IMAGE:2900313, mRNA, complete cds.	667	88
1525	gi|12653255	Homo sapiens	ubiquitin-conjugating enzyme E2N (homologous to yeast UBC13), clone MGC:8489 IMAGE:2822013, mRNA, complete cds.	667	88
1526	AAY87271	Homo sapiens	INCY- Human signal peptide containing protein HSPP-48 SEQ ID NO:48.	471	86
1526	gi|17128086	Corynebacterium glutamicum	cdsA	70	27
1526	gi|12544226	Corynebacterium glutamicum	RXA01894	70	27
1527	gi|13358942	Macaca fascicularis	hypothetical protein	2660	97
1527	AAB87587	Homo sapiens	GETH Human PRO1693.	2647	100
1527	AAU12439	Homo sapiens	GETH Human PRO1693 polypeptide sequence.	2647	100
1528	AAB425	Homo sapiens	CURA- Human ORFX	2137	99

SEQ ID	Hit ID	Speicies	Description	S score	Percent identity
	73		ORF2337 polypeptide sequence SEQ ID NO:4674.		
1528	gi12845823	Mus musculus	putative	1792	95
1528	gi4495063	Homo sapiens	Human DNA sequence from clone 108K11 on chromosome 6p21 Contains SRP20 (SR protein family member), Ndr protein kinase gene similar to yeast suppressor protein SRP40, EST and GSS, complete sequence.	1468	99
1529	gi15823636	Homo sapiens	ALS2 mRNA, complete cds, long form.	8660	99
1529	gi16076812	Homo sapiens	alsin mRNA, complete cds.	8646	99
1529	gi15823640	Mus musculus	Als2	8005	91
1530	AAG64171	Homo sapiens	TAKE Human profilin IIL.	750	100
1530	gi12804213	Homo sapiens	profilin 2, clone MGC:1684 IMAGE:3533907, mRNA, complete cds.	750	100
1530	gi10952520	Homo sapiens	profilin IIa (PFN2) mRNA, complete cds, alternatively spliced.	750	100
1531	AAG64171	Homo sapiens	TAKE Human profilin IIL.	636	87
1531	gi12804213	Homo sapiens	profilin 2, clone MGC:1684 IMAGE:3533907, mRNA, complete cds.	636	87
1531	gi10952520	Homo sapiens	profilin IIa (PFN2) mRNA, complete cds, alternatively spliced.	636	87
1532	AAB94952	Homo sapiens	HELI- Human protein sequence SEQ ID NO:16482.	1830	99
1532	gi17907791	Homo sapiens	TAIP-2 mRNA for TGF-beta induced apotosis protein 2, complete cds.	1830	99
1532	gi10433016	Homo sapiens	cDNA FLJ11703 fis, clone HEMBA1005075.	1830	99
1533	gi14133223	Homo sapiens	mRNA for KIAA0876 protein, partial cds.	4559	100
1533	gi6910563	Homo sapiens	chromosome 19, BC335474 (CIT-HSPC_482H14), complete sequence.	4370	99
1533	gi13938056	Mus musculus	Similar to KIAA0677 gene product	3313	73
1534	gi7243191	Homo sapiens	mRNA for KIAA1405 protein, partial cds.	3986	99
1534	gi12313647	Mus musculus	MmKIF17	3319	77
1534	gi4102184	Homo sapiens	KIF3-related motor protein (KIF3X) mRNA, partial cds.	1084	90
1535	gi12053311	Homo sapiens	mRNA; cDNA DKFZp434K229 (from clone DKFZp434K229); complete cds.	1600	100